

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 10:14:14 : Search time 9019.27 Seconds
(without alignments)
11121.789 Million cell updates/sec

Title: US-09-945-326-1
Perfect score: 2452
Sequence: 1 cgtgtgtgtctcctgcggc.....actgttaaaaaaaaaa 2452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. NO. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2452	100.0	2452	6	AX394747 Sequence
2	2445.8	99.7	2494	9	BC007970 Homo sapi
3	2422.4	98.8	2448	9	BC013354 Homo sapi
4	2391.8	97.5	2411	9	AF327351 Homo sapi
5	2382.2	97.2	2399	6	BD158773 Homo sapi
6	2382.2	97.2	2399	9	BD158773 Primer fo
7	2115.4	86.3	2400	6	AK024012 Homo sapi
8	2115.4	86.3	2400	6	BD157796 Primer fo
9	1863	76.0	1863	6	AK022568 Homo sapi
10	1439	58.7	1494	9	AX394749 Sequence
11	1388.4	56.6	2129	10	BC001817 Homo sapi
12	1388.4	56.6	2148	10	BC033277 Mus muscu
13	1388.4	56.6	3071	10	BC032213 Mus muscu
14	759.6	31.0	1995	9	BC031137 Mus muscu
15	750.6	30.6	771	6	AK027186 Homo sapi
16	608	24.8	17777	2	BD150623 Primer fo
17	608	24.8	185470	2	AC022993 Homo sapi
18	530.4	21.6	544	6	AC112484 Homo sapi
19	518.4	21.1	543	6	BD109621 EST and e
20	499.6	20.4	772	6	AX588711 Sequence
21	487.8	19.9	563	6	BD129328 Human gen
22	487	19.9	507	6	BD155541 Primer fo
23	439.2	17.9	796	6	AX588841 Sequence
24	418.4	17.1	431	6	BD147321 Primer fo
25	396.4	16.2	2058	9	AX072562 Sequence
26	396.4	16.2	2103	10	AB083302 Macaca fa
27	396	16.2	2140	4	BTU30817 Mus muscu
28	394.8	16.1	2171	10	U30817 Bos taurus
29	391.6	16.0	1965	6	BC026559 Mus muscu
30	391.6	16.0	2168	9	BD141736 IRAP bind
31	391.6	16.0	2169	9	BC020218 Homo sapi
32	391.6	16.0	2219	9	D43682 Human mRNA
33	391.6	16.0	2231	9	HSHVLCAD
34	391.6	16.0	2258	9	BC012912 Homo sapi
35	391.6	16.0	2259	9	AK056083 Homo sapi
36	391.2	16.0	1840	10	BC000399 Homo sapi
37	390.4	15.9	1439	10	AF017176 Mus muscu
38	388.6	15.8	2117	10	AY292988 Rattus no
39	388.6	15.8	2117	10	AX401908 Sequence
40	380.2	15.5	434	6	D30647 Rat mRNA fo
41	372.8	15.2	171981	3	AX071059 Sequence
42	372.8	15.2	176813	3	AC008344 Drosophill
43	372.8	15.2	176980	2	AC008096 Drosophill
44	372.8	15.2	277421	3	AC020290 Drosophill
45	372.6	15.2	412	6	AE003796 Drosophill
					AX071159 Sequence

ALIGNMENTS

RESULT	1	AX394747	Sequence 1 from Patent WO0218582.	2452 bp	DNA	linear	PAT 18-MAY-2002
LOCUS	AX394747	Sequence 1 from Patent WO0218582.					
DEFINITION	AX394747						
ACCESSION	AX394747						
VERSION	AX394747.1	GI:21065826					
KEYWORDS		Homo sapiens (human)					
SOURCE		Homo sapiens					
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.					
REFERENCE	1	Meyers, R. and Hunter, J.J.					
AUTHORS		62112, a novel human dehydrogenase and uses thereof					
TITLE		Patent: WO 0218582-A 1 07-MAR-2002;					
JOURNAL							

FEATURES		Millennium Pharmaceuticals, Inc. (US)
Source	Location/Qualifiers,	
	1..2452	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	67..1932	
CDS	note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD32049.1"	
	/db_xref="GI:21065827"	
	/translation="MSGGGLFRTTAAARACGLVSTANRLRLTSPVRAFAKELF	
	LGKIKKVEFPPEVSDDELNEINOF LGPVEKFEFEVDSDKIDQEGRPIDTEFLK	
	SLGLFGLQPEYGLGFSNTMYSRIGETIISMDGSIITVLAHQAIGLKGIIILACTEE	
	QKAYLPKLASGEHIAAFCLTEPASGSDAASIRSRATLSEDKHVLNLSKVWVINGG	
	LANIFTFAKTEVDSDGSKDKITAFIVERDFGGVINGKPEDKLIGIRSNTEVHF	
	NKIPVNTILGVDGFKVAMNLSNRSFSGMSVAGLLRLDIEMTAETAYCTKGFNK	
	RLSEFLGIEKALMAQKAYVMESMTYLTAGMLDQFPDCKSIEAAMYVSSSEAAWQ	
	CYSEALQILGGIYTRDPYERILRDTRLILLFEGTNEILRMVIALTLQIHAGRIIT	
	RIHELKQAKSVMDTVGRRLRDSLGRVLDLGTNGHVHPSLADLSANKFEENTYCF	
	GRTEVTLRLRFKTIIMEQQLVLRVANILINLXGMTAVLSRASRSIRIGLRNHDHVL	
	LANTFCEAYLQNLFSLSQDAIPENLDEQIKKVSQQLLEKRAYICAHPLDRTC"	
BASE COUNT	616 a 616 c 685 g 535 t	
ORIGIN		
Query Match		100.0%; Score 2452; DB 6; Length 2452;
Best Local Similarity		100.0%; Pred No. 0;
Matches 2452; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CGTGTGTGTCTCCCTCGCGCGCTAAGAAGGGAGAGCTAGGCTGAGGCTGGGGAAACATCG 60
Db	1	CGTGTGTGTCTCCCTCGCGCGCTAAGAAGGGAGAGCTAGGCTGAGGCTGGGGAAACATCG 60
Qy	61	GGCAGCATGAGCGCTCGCGGCTCTCCCTCGCACACAGCGCTCGCGCTCGTGCCTGCCGG 120
Db	61	GGCAGCATGAGCGCTCGCGGCTCTCCCTCGCACACAGCGCTCGCGCTCGTGCCTGCCGG 120
Qy	121	GGTCTGTGTGTCTACCGGGAACCGCGGCTACTCGCGACACAGCGCTCGCGCTCGTGCCTGCCGG 180
Db	121	GGTCTGTGTGTCTACCGGGAACCGCGGCTACTCGCGACACAGCGCTCGCGCTCGTGCCTGCCGG 180
Qy	181	TTCCGCCAAGAGCTTTTCTTAGCAAAATCAAGAAGAAAGAGTTTCCCATTTCCAGAA 240
Db	181	TTCCGCCAAGAGCTTTTCTTAGCAAAATCAAGAAGAAAGAGTTTCCCATTTCCAGAA 240
Qy	241	GTTAGCCAAGATGAACATTAATGAATCAATCAGTTCTTGGACCCCGTGGAAAAATCTTC 300
Db	241	GTTAGCCAAGATGAACATTAATGAATCAATCAGTTCTTGGACCCCGTGGAAAAATCTTC 300
Qy	301	ACTGAAGAGTGGACTCCCGAAAAATTGACAGGAAGGAAATCCACAGATGAACATTTG 360
Db	301	ACTGAAGAGTGGACTCCCGAAAAATTGACAGGAAGGAAATCCACAGATGAACATTTG 360
Qy	361	GAGAAATGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAAAGAAATATGTTGGCGTG 420
Db	361	GAGAAATGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAAAGAAATATGTTGGCGTG 420
Qy	421	GGCTTCTCCACACCATGACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATC 480
Db	421	GGCTTCTCCACACCATGACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATC 480
Qy	481	ACTGTGACCTGGACGCCACAGGCTATTGGCTCAAGGGGATCATCTGGCTGGCACT 540
Db	481	ACTGTGACCTGGACGCCACAGGCTATTGGCTCAAGGGGATCATCTGGCTGGCACT 540
Qy	541	GAGGAGCAGAAAGCCAAATCTTGCTTAACTGGCTCCGGGGAGCAATTCAGAGCTTC 600
Db	541	GAGGAGCAGAAAGCCAAATCTTGCTTAACTGGCTCCGGGGAGCAATTCAGAGCTTC 600
Qy	601	TGCCTCAGGAGCCAGCAGTGGGAGCGATGCAGCCTCAATCCGAGCAGAGCCACACTA 660
Db	601	TGCCTCAGGAGCCAGCAGTGGGAGCGATGCAGCCTCAATCCGAGCAGAGCCACACTA 660

Qy	661	AGTGAAGACAAAGAGCACTACATCTCTCAATGGCTCCAAAGCTCTGGATTACTAATGGAGGA 720
Db	661	AGTGAAGACAAAGAGCACTACATCTCTCAATGGCTCCAAAGCTCTGGATTACTAATGGAGGA 720
Qy	721	CTGGCCCAATATTTTACTGTGTTTGAAGACTGAGGTCTGTTGATCTGTGATGATCAGT 780
Db	721	CTGGCCCAATATTTTACTGTGTTTGAAGACTGAGGTCTGTTGATCTGTGATGATCAGT 780
Qy	781	AAAGACAAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCATCTAATGGGAAA 840
Db	781	AAAGACAAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCATCTAATGGGAAA 840
Qy	841	CCCAGAAATAATTTAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACAACACC 900
Db	841	CCCAGAAATAATTTAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACAACACC 900
Qy	901	AAGATACCTGTGGAATAACATCTTGGAGAGTCGAGATGGGTTTAAAGTGGCCATGAAC 960
Db	901	AAGATACCTGTGGAATAACATCTTGGAGAGTCGAGATGGGTTTAAAGTGGCCATGAAC 960
Qy	961	ATCCTCAACAGCGCGGCTTCCAGCATGGGCGCTGCTGGCTGGGCTGCTCAAGAGATTG 1020
Db	961	ATCCTCAACAGCGCGGCTTCCAGCATGGGCGCTGCTGGCTGGGCTGCTCAAGAGATTG 1020
Qy	1021	ATTGAATGACTGTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGCTCAGTGAA 1080
Db	1021	ATTGAATGACTGTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGCTCAGTGAA 1080
Qy	1081	TTTGGATTGATTACAGGAGAAATTTGCATGTGCTCAGAGGCTTAGCTCATGGAGAGT 1140
Db	1081	TTTGGATTGATTACAGGAGAAATTTGCATGTGCTCAGAGGCTTAGCTCATGGAGAGT 1140
Qy	1141	ATGACCTACCTCACAGCAGGGATGCTGCACCAACCTGCTTCCGACTGCTCCATCCAG 1200
Db	1141	ATGACCTACCTCACAGCAGGGATGCTGCACCAACCTGCTTCCGACTGCTCCATCCAG 1200
Qy	1201	GCAGCCATGGTGAAGGTGTTTACGCTCCGAGGCGGCTGGCAGTGTGTGAGTGAGCGCTG 1260
Db	1201	GCAGCCATGGTGAAGGTGTTTACGCTCCGAGGCGGCTGGCAGTGTGTGAGTGAGCGCTG 1260
Qy	1261	CAGATCCTCGGGGCTTGGGCTACACAGGGACTATCCGTACGAGCGGATCTCGCTGAC 1320
Db	1261	CAGATCCTCGGGGCTTGGGCTACACAGGGACTATCCGTACGAGCGGATCTCGCTGAC 1320
Qy	1321	ACCGGATCCTCTCTATCTTCGAGGGAACCAATGAGATTCTCCGGATGATGATGCGCTG 1380
Db	1321	ACCGGATCCTCTCTATCTTCGAGGGAACCAATGAGATTCTCCGGATGATGATGCGCTG 1380
Qy	1381	ACGGGTCTGCAGATGCCGGCGCATCTCTGACTACAGGATCCATGAGCTTAAACAGGCC 1440
Db	1381	ACGGGTCTGCAGATGCCGGCGCATCTCTGACTACAGGATCCATGAGCTTAAACAGGCC 1440
Qy	1441	AAAGTCAGCAGCTCATGATACCGTTGGCGGAGGCTTCCGGACTTCCCTGGGCGCAACT 1500
Db	1441	AAAGTCAGCAGCTCATGATACCGTTGGCGGAGGCTTCCGGACTTCCCTGGGCGCAACT 1500
Qy	1501	GTGACCTGGGCTGACAGCAACCATGAGTTGTGACCCAGTCTTGGGACAGTGC 1560
Db	1501	GTGACCTGGGCTGACAGCAACCATGAGTTGTGACCCAGTCTTGGGACAGTGC 1560
Qy	1561	AACAAGTTTGAAGAGAACACCTTACTGCTTCGGCGGACCGTGGAGACACTGTGCTCCGC 1620
Db	1561	AACAAGTTTGAAGAGAACACCTTACTGCTTCGGCGGACCGTGGAGACACTGTGCTCCGC 1620
Qy	1621	TTTGGCAAGACCATCATGGAGGAGCAGTGGTACTGAAGCGGTGGCCAACTCTCTCATC 1680
Db	1621	TTTGGCAAGACCATCATGGAGGAGCAGTGGTACTGAAGCGGTGGCCAACTCTCTCATC 1680
Qy	1681	AACCTGTATGGATGAGCGCGCTGCTCGCGGCGGACCGCTCCATCCGCTTGGGCTC 1740
Db	1681	AACCTGTATGGATGAGCGCGCTGCTCGCGGCGGACCGCTCCATCCGCTTGGGCTC 1740
Qy	1741	CGCAACCCAGCAGGAGTCTCTTGGCCAAACACCTTCTTSCGTGGGAAGCTTACTTGCAG 1800

Tue Sep 9 10:14:35 2003

GRVETILLREFKTIMEQVLKRVANILINLYGMTAVLSRARSIRIGLRNHDHVL
LANTFVAYLQNLFSLSOLDKYAPENLDEQIKKVSQOILEKRAYICAHPLDRTC

BASE COUNT	643 a	621 c	692 g	538 t	
ORIGIN					
Query Match	99.7%	Score 2445.8	DB 9	Length 2494	
Best Local Similarity	99.9%	Pred. No. 0			
Mismatches 2447	Conservative	0	Mismatches 2	Indels 0	Gaps 0
QY	4	GTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGCTGAGGCTGGGGGACATCGGGC	63		
DB	26	GTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGCTGAGGCTGGGGGACATCGGGC	85		
QY	64	AGCATGAGGGGTGCGGGCTCTTCCTGCGCACACCGGTGCGGTCTGTCCTGCGGGGT	123		
DB	86	AGCATGAGGGGTGCGGGCTCTTCCTGCGCACACCGGTGCGGTCTGTCCTGCGGGGT	145		
QY	124	CTGGTGTCTTACCGCGAACCGGGCTTACTGCGCACACCGCGCTGTACGAGCTTTC	183		
DB	146	CTGGTGTCTTACCGCGAACCGGGCTTACTGCGCACACCGCGCTGTACGAGCTTTC	205		
QY	184	GCCAAAGAGCTTTTCTAGSCAAAATCAAGAAAGAAAGATTTTCCCATTTCCAGAGTT	243		
DB	206	GCCAAAGAGCTTTTCTAGSCAAAATCAAGAAAGAAAGATTTTCCCATTTCCAGAGTT	265		
QY	244	AGCCAGATGAATTAATGAATCAATCAGTTCTTGGGACCGGTGGAAAAATTTCTTCACT	303		
DB	266	AGCCAGATGAATTAATGAATCAATCAGTTCTTGGGACCGGTGGAAAAATTTCTTCACT	325		
QY	304	GAAGAGTGGACTCCCGAAAAATTTGACAGGAAGGAAAAATCCAGATGAAGTTGGAG	363		
DB	326	GAAGAGTGGACTCCCGAAAAATTTGACAGGAAGGAAAAATCCAGATGAAGTTGGAG	385		
QY	364	AAATGAAGAGCTTAGGGCTTTTGGGCTCAAGTCCCAAGAAATATGTTGGCTGGGC	423		
DB	386	AAATGAAGAGCTTAGGGCTTTTGGGCTCAAGTCCCAAGAAATATGTTGGCTGGGC	445		
QY	424	TTCTCCAAACCATGTACTCAAGACTAGGGAGATCATCAGATGATGGTCCATCACT	483		
DB	446	TTCTCCAAACCATGTACTCAAGACTAGGGAGATCATCAGATGATGGTCCATCACT	505		
QY	484	GTGACCTGGCAGCGCACCAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCAGTGA	543		
DB	506	GTGACCTGGCAGCGCACCAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCAGTGA	565		
QY	544	GAGCAGAAGCCAAATACTTCCCTAAACTGSCGTCCCGGGAGACATTTGAGCCTTCTGC	603		
DB	566	GAGCAGAAGCCAAATACTTCCCTAAACTGSCGTCCCGGGAGACATTTGAGCCTTCTGC	625		
QY	604	CTCAGGAGCCAGCCAGTGGGAGGATCGAGCCTCAATCCGAGCAGAGCCACACTAAGT	663		
DB	626	CTCAGGAGCCAGCCAGTGGGAGGATCGAGCCTCAATCCGAGCAGAGCCACACTAAGT	685		
QY	664	GAAGCAAGAAGCACTACATCTCAATGGCTCCAGGCTCTGGATTACTAATGGAGGACTG	723		
DB	686	GAAGCAAGAAGCACTACATCTCAATGGCTCCAGGCTCTGGATTACTAATGGAGGACTG	745		
QY	724	GCCAAATTTTACTGTGTTTCCAAAGACTGAGGTCTTGTATCTTGATGGATCAGTGAAA	783		
DB	746	GCCAAATTTTACTGTGTTTCCAAAGACTGAGGTCTTGTATCTTGATGGATCAGTGAAA	805		
QY	784	GACAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAACCC	843		
DB	806	GACAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAACCC	865		
QY	844	GAAGATAAATAGGCAATTCGGGGTCCCAACACTTGTGAGTCCATTTTGAAGAACCAAG	903		
DB	866	GAAGATAAATAGGCAATTCGGGGTCCCAACACTTGTGAGTCCATTTTGAAGAACCAAG	925		
QY	904	ATACCTGTGGAAGCAATCTTGGAGAGTGGGAGTGGGTTTAAAGTGGCCATGAACATC	963		
DB	926	ATACCTGTGGAAGCAATCTTGGAGAGTGGGAGTGGGTTTAAAGTGGCCATGAACATC	985		

QY	964	CTCAACAGCGCGCGTTCAGCATGGCAGCGCTCGTGGCTGCTCAAGAGATTGATT	1023
DB	986	CTCAACAGCGCGCGTTCAGCATGGCAGCGCTCGTGGCTGCTCAAGAGATTGATT	1045
QY	1024	GAATCACTGCTCAGTACCGCTGCACAAGAAACAGTTTAAACAAGAGGCTCAGTGAATTT	1083
DB	1046	GAATCACTGCTCAGTACCGCTGCACAAGAAACAGTTTAAACAAGAGGCTCAGTGAATTT	1105
QY	1084	GGATTGATTCAAGAGAAATTTGCACTGATGGCTCAGAAGGCTTACGTATGAGAGTAGT	1143
DB	1106	GGATTGATTCAAGAGAAATTTGCACTGATGGCTCAGAAGGCTTACGTATGAGAGTAGT	1165
QY	1144	ACCTACCTCACAGAGGATGCTGGACCAACCTGGCTTCCCGACTGCTCCATCGAGGCA	1203
DB	1166	ACCTACCTCACAGAGGATGCTGGACCAACCTGGCTTCCCGACTGCTCCATCGAGGCA	1225
QY	1204	GCCATGTGAAGTGTTCAGCTCCGAGCGCGCTGGCAGTGTGTAGTGAGGCGCTGCAG	1263
DB	1226	GCCATGTGAAGTGTTCAGCTCCGAGCGCGCTGGCAGTGTGTAGTGAGGCGCTGCAG	1285
QY	1264	ATCCTCGGGGCTTGGGCTACACAAGGAGTATCCGTACGAGCGCATACTGCGTGACACC	1323
DB	1286	ATCCTCGGGGCTTGGGCTACACAAGGAGTATCCGTACGAGCGCATACTGCGTGACACC	1345
QY	1324	CGCATCTCTCATCTTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCGCTGACG	1383
DB	1346	CGCATCTCTCATCTTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCGCTGACG	1405
QY	1384	GGTCTGACAGTATGCGCGCGCATCTGACTACCAAGATCCATGAGCTTAAACAGGCCAAA	1443
DB	1406	GGTCTGACAGTATGCGCGCGCATCTGACTACCAAGATCCATGAGCTTAAACAGGCCAAA	1465
QY	1444	GTGACACAGTCATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCACTGTG	1503
DB	1466	GTGACACAGTCATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCACTGTG	1525
QY	1504	GACCTGGGCTGACAGGCAACCATGAGTTGTGACCCCTCTTGGGAGACACTGCTCCTCAAC	1563
DB	1526	GACCTGGGCTGACAGGCAACCATGAGTTGTGACCCCTCTTGGGAGACACTGCTCCTCAAC	1585
QY	1564	AGTTTGGAGGAACACACTACTGCTTCGCGCGGAGCCGCTGGAGACACTGCTCCTCAAC	1623
DB	1586	AGTTTGGAGGAACACACTACTGCTTCGCGCGGAGCCGCTGGAGACACTGCTCCTCAAC	1645
QY	1624	GGCAAGACCATCATGGAGAGCAGTGGTACTGAAGCGGTGGCCAACTCCTCATCAAC	1683
DB	1646	GGCAAGACCATCATGGAGAGCAGTGGTACTGAAGCGGTGGCCAACTCCTCATCAAC	1705
QY	1684	CTGTATGGCATGACGCGCTGCTGTCGGGCGGAGCCGCTCCATCCGATTTGGGCTCCGC	1743
DB	1706	CTGTATGGCATGACGCGCTGCTGTCGGGCGGAGCCGCTCCATCCGATTTGGGCTCCGC	1765
QY	1744	AACCAAGCAGCAGAGGTTCTTTGGCCAACTTCTGCGTGAAGCTTACTTTGAGAGT	1803
DB	1766	AACCAAGCAGCAGAGGTTCTTTGGCCAACTTCTGCGTGAAGCTTACTTTGAGAGT	1825
QY	1804	CTCTTCAGCTCTCTCAGCTGGCAAGATGCTTCCAGAAACCTAGATGAGCAGATTAG	1863
DB	1826	CTCTTCAGCTCTCTCAGCTGGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAG	1885
QY	1864	AAAGTGTCCAGCAGATTCCTTTGAGAGCGAGCCTATATCTGTGCCACCCCTCTGGACGG	1923
DB	1886	AAAGTGTCCAGCAGATTCCTTTGAGAGCGAGCCTATATCTGTGCCACCCCTCTGGACGG	1945
QY	1924	ACATGCTGAGCGAGGAGCAGTGTCCCTGCTACCCGCGCCCTTACCCTATGGCCGCTTG	1983
DB	1946	ACATGCTGAGCGAGGAGCAGTGTCCCTGCTACCCGCGCCCTTACCCTATGGCCGCTTG	2005
QY	1984	CTGGATGACTGTTACTCTTTTTCAGAGGTGTTGGATTATCACAGGTTAAGCCTTTTG	2043
DB	2006	CTGGATGACTGTTACTCTTTTTCAGAGGTGTTGGATTATCACAGGTTAAGCCTTTTG	2065
QY	2044	TTCCCCGCTGTCACCTGAAGGTTGTGCGCTGGGCTGGGAGAGCCTCTTCCAGGTTTGA	2103


```

Db      2066  TTCCCGCTCTCACCTGAAGGTTGTGCGCTGGGAGAGCTCTTCCAGGTTTGA 2125
Qy      2104  CCTCAGGACGCTCTCTTACAGGACCATCACAGCTTCTGAAGTGGGAGAGAGAG 2163
Db      2126  CCTCAGGACGCTCTCTTACAGGACCATCACAGCTTCTGAAGTGGGAGAGAGAG 2185
Qy      2164  AATGGAATTCGTGACCTGGAAGTGGGAGTATTCGTCATTGAGGACACCATAGT 2223
Db      2186  AATGGAATTCGTGACCTGGAAGTGGGAGTATTCGTCATTGAGGACACCATAGT 2245
Qy      2224  GGAACCTGGGCTTATGCTGCTCCAGGCTGAGTGGTGGGAGCTGTGCAGG 2283
Db      2246  GGAACCTGGGCTTATGCTGCTCCAGGCTGAGTGGTGGGAGCTGTGCAGG 2305
Qy      2284  TGTGGATGAGCATTTCTGCTCAACACACATTCCTTAAGAACACGCTTGAAGCTGTGC 2343
Db      2306  TGTGGATGAGCATTTCTGCTCAACACACATTCCTTAAGAACACGCTTGAAGCTGTGC 2365
Qy      2344  TGGGTCTATTCATTAACATAGAGAGAGGACCTTAAACATGTACAGGACCATTTAA 2403
Db      2366  TGGGTCTATTCATTAACATAGAGAGAGGACCTTAAACATGTACAGGACCATTTAA 2425
Qy      2404  CAAAGATATTAATGTCAATCTGTACTGTCTTAAAAAATAAAAAA 2452
Db      2426  CAAAGATATTAATGTCAATCTGTACTGTCTTAAAAAATAAAAAA 2474

```

RESULT 3

LOCUS

BC013354 2448 bp mRNA linear PRI 20-MAY-2003
 Homo sapiens very-long-chain acyl-CoA dehydrogenase-VLCAD, mRNA
 (CDNA clone MGC:14970 IMAGE:3935925), complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BC013354.1 GI:15426497
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Strausberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Brat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitling, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalinski, D.E.,
 Schnerker, A., Schein, J.E., Jones, S.J., and Warra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 2448)
 Strausberg, R.
 Direct Submission
 Submitted (31-AUG-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

Email: cgabbs-re@mail.nih.gov
 Tissue procurement: DCTD/Dnp
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amadansystemsbio.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 25 Row: c Column: 21.

FEATURES

source

1..2448
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:14970 IMAGE:3935925"
 /tissue_type="Lung, small cell carcinoma"
 /clone_lib="NIH_MGC_7"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..2448
 /gene="VLCAD"
 /note="synonyms: NP0002, MGC14452, ACAD-9"
 /db_xref="LocusID:28976"
 66..1931
 /codon_start=1
 /product="very-long-chain acyl-CoA dehydrogenase VLCAD"
 /protein_id="AAH13354.1"
 /db_xref="GI:15426498"
 /translation="MSCGELFLRTTAAARGLVSTANRLRLTSPVPVAFKELF
 LGRKKKEVPEVPEVQDELNEINQFLGPVEKFTFEVDSRKIDQSGKIPDETLEKLK
 SLGLQVLPVEEYGLGNTMYSLRGLISDGSITVTLAAHQALGLGILAGTLEE
 OKAKYLPKLAGSHEIAAFCLTEPASGSDAASIRSRATLSEDKHLYLNGSKVWITNGG
 LANITVEAKTEVYVSDGSKITAFIVERDGGVTNGKPKDLGINSKVTNVE
 NTKIPVENILGEVGGFKVAMNINLRSFMSVAGLLKRLIETMAVYKRFQFNK
 CVSEALQIGLGLYTRDYPYRILRLTRILLIPEGTNEILRMVIALTQHQHRLIT
 RIHELQAKYSTVMDTVGRRLRDLGTLGTHGVVHPSLDSANKFRENITYCF
 GRVTEVLLRFKGTIMEEQILKRVANILNYGMTAVLSRASRIRIGLRNHDVEVL
 LANTFCEVAYLQNLFSQLDKYAPENLDEQIKKVSQQLLEKRAYICAHPLDRTC"

gene

BASE COUNT 624 a 617 c 682 g 525 t
 ORIGIN
 Query Match 98.8%; Score 2422.4; DB 9; Length 2448;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 2429; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

CDS

10 GTCCCTGCGGCGCTAAGAGGGGAGACTGAGCTGAGGCTGGGGAACATCGGGCAGCATG 69
 9 GTCCCTGCGGCGCTAAGAGGGGAGACTGAGCTGAGGCTGGGGAACATCGGGCAGCATG 68
 70 AGCGGTGCGGCGCTCTTCTCGGCACACAGCGCTGCGGCTCGCTGCGGCGGTCTGGTG 129
 69 AGCGGTGCGGCGCTCTTCTCGGCACACAGCGCTGCGGCTCGCTGCGGCGGTCTGGTG 128
 130 GTCTCTACCGCAACCGGCGGCTACTGGCGCACCGCGGCTGTACGAGCTTTCGCCAAA 189
 129 GTCTCTACCGCAACCGGCGGCTACTGGCGCACCGCGGCTGTACGAGCTTTCGCCAAA 188
 190 GAGCTTTTCTAGGCAAAATCAAGAAGAAAGTTTCCCATTTCCAGAGTTAGCCAA 249
 189 GAGCTTTTCTAGGCAAAATCAAGAAGAAAGTTTCCCATTTCCAGAGTTAGCCAA 248
 250 GATGAACCTAATGAATCAATCAGTCTTGGGACCGCTGGAAATTTCTTCACTGAAGAG 309
 249 GATGAACCTAATGAATCAATCAGTCTTGGGACCGCTGGAAATTTCTTCACTGAAGAG 308
 310 GTGGACTCCGCAAAATTTGACGAGAGGAAATCCCGATGAAACTTTGAGAAATTTG 369

Db	309	GTGGACTCCGAAAAATTGACAGGAGGAAAAATCCAGATGAAACTTTGGAGAAATTG	368	1450	ACAGTCATGATACCCGTTGGCCGAGAGGCTTCCGGAATCTCCCTGGCCGAACTGTGGACCTG	1509
Qy	370	AAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGCTTCNC	429	1449	ACAGTCATGATACCCGTTGGCCGAGAGGCTTCCGGAATCTCCCTGGCCGAACTGTGGACCTG	1508
Db	369	AAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGCTTCNC	428	1510	GGGCTGACAGGCAACCATGAGTTGTGCACCCAGTCTTCCGACAGTGCACCAAGTTT	1569
Qy	430	AACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTGACC	489	1509	GGGCTGACAGGCAACCATGAGTTGTGCACCCAGTCTTCCGACAGTGCACCAAGTTT	1568
Db	429	AACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTGACC	488	1570	GAGGAGAACACTACTGCTTCCGCGGACCGTGGAGACACTGTGTCCGCTTTGGCAAG	1629
Qy	490	CTGGCAGCGCACCAGGCTATTGGCTTCAAGGGGATCATCTTGGCTGGCACTGAGAGGAG	549	1569	GAGGAGAACACTACTGCTTCCGCGGACCGTGGAGACACTGTGTCCGCTTTGGCAAG	1628
Db	489	CTGGCAGCGCACCAGGCTATTGGCTTCAAGGGGATCATCTTGGCTGGCACTGAGAGGAG	548	1630	ACCATCATGAGGAGGAGTGGTACTGAAAGGGGTGGCCAACTCCCTCATCAACCTGTAT	1689
Qy	550	AAAGCCAAATACTTGGCTTAACTGGCGTCCGGGGAGCACATTTGAGGCTTCTGCCTCAGG	609	1629	ACCATCATGAGGAGGAGTGGTACTGAAAGGGGTGGCCAACTCCCTCATCAACCTGTAT	1688
Db	549	AAAGCCAAATACTTGGCTTAACTGGCGTCCGGGGAGCACATTTGAGGCTTCTGCCTCAGG	608	1690	GGCATGACGGCGTGTCTCCGCGGACCGTCCATCCGCAATGGGCTCCGCAACAC	1749
Qy	610	GAGCCAGCAGTGGGAGGATGAGGCTCAATCCGAGGAGAGCCACACTAAGTGAAGAC	669	1689	GGCATGACGGCGTGTCTCCGCGGACCGTCCATCCGCAATGGGCTCCGCAACAC	1748
Db	609	GAGCCAGCAGTGGGAGGATGAGGCTCAATCCGAGGAGAGCCACACTAAGTGAAGAC	668	1750	GACCAAGGAGTGTCTTGGCCAAACACTTCTGGTGGAGGTTACTTGCAGAACTCTTC	1809
Qy	670	AAGAGCAGTACACTTCAATGGCTCCAGGCTCGATTAATGAGGAGTGGCCAT	729	1749	GACCAAGGAGTGTCTTGGCCAAACACTTCTGGTGGAGGTTACTTGCAGAACTCTTC	1808
Db	669	AAGAGCAGTACACTTCAATGGCTCCAGGCTCGATTAATGAGGAGTGGCCAT	728	1810	AGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTTAGATGAGCAGATTAAAGAAATG	1869
Qy	730	ATTTTACTGTGTTGCAAAAGACTGAGGCTGATTGATTCATGATGATGATGAAAGACAA	789	1809	AGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTTAGATGAGCAGATTAAAGAAATG	1868
Db	729	ATTTTACTGTGTTGCAAAAGACTGAGGCTGATTGATTCATGATGATGATGAAAGACAA	788	1870	TCCCAGCAGATCTTTGAGAAGGAGCCATATCTGTGCCACCCTCTGGACAGGACATGC	1929
Qy	790	ATCAGCAGTATGATGAGGAGTGGTGGAGTCACTAATGGGAAACCCGAGAT	849	1869	TCCCAGCAGATCTTTGAGAAGGAGCCATATCTGTGCCACCCTCTGGACAGGACATGC	1928
Db	789	ATCAGCAGTATGATGAGGAGTGGTGGAGTCACTAATGGGAAACCCGAGAT	848	1930	TGAGCAGGAGGAGTGTCCCTGTACCGCCGCCCTACCCATGGCCGCTTGTGTGAT	1989
Qy	850	AAATAGGCAATCGGGCTCCACACTGTGAGTCCATTTGAAACACCAAGATACCT	909	1929	TGAGCAGGAGGAGTGTCCCTGTACCGCCGCCCTACCCATGGCCGCTTGTGTGAT	1988
Db	849	AAATAGGCAATCGGGCTCCACACTGTGAGTCCATTTGAAACACCAAGATACCT	908	1990	GACTGTACTCTTTTTCAGAAAGGTTGGGATTTATCACAGGTTAAGCTTTTGTTCCTC	2049
Qy	910	GTGAAAAACATCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACCTCAAC	969	1989	GACTGTACTCTTTTTCAGAAAGGTTGGGATTTATCACAGGTTAAGCTTTTGTTCCTC	2048
Db	909	GTGAAAAACATCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACCTCAAC	968	2050	GTCTGCACCTGAAGGTTGTCCGCTGGCCGAGAGGCTCTTCCAGGTTTGTGACCTGCA	2109
Qy	970	ACCGCCCGTTCAGCATGGCAGGCTGTGCTGGCTGCTCAAGAGATTGATTGAATG	1029	2049	GTCTGCACCTGAAGGTTGTCCGCTGGCCGAGAGGCTCTTCCAGGTTTGTGACCTGCA	2108
Db	969	ACCGCCCGTTCAGCATGGCAGGCTGTGCTGGCTGCTCAAGAGATTGATTGAATG	1028	2110	GGCAGTGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCCGAGAGAGAGATGGA	2169
Qy	1030	ACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTG	1089	2109	GGCAGTGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCCGAGAGAGAGATGGA	2168
Db	1029	ACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTG	1088	2170	ATTGCTGACCCCTGGAACTGGCGGGTATTCTGGTCATTGAGGAGACCATAGTGGAAAC	2229
Qy	1090	ATTGAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATGACCTAC	1149	2169	ATTGCTGACCCCTGGAACTGGCGGGTATTCTGGTCATTGAGGAGACCATAGTGGAAAC	2228
Db	1089	ATTGAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATGACCTAC	1148	2230	TGGGGCTTATGCTGCTCCCTCCAGGCTGTGAGTGGGTGGGACCTGTGTGAGGTGGA	2289
Qy	1150	CTCAGAGGAGGATCTGACCAACCTGGCTTCCCGACTGCTCCATCGAGGAGCCCATG	1209	2229	TGGGGCTTATGCTGCTCCCTCCAGGCTGTGAGTGGGTGGGACCTGTGTGAGGTGGA	2288
Db	1149	CTCAGAGGAGGATCTGACCAACCTGGCTTCCCGACTGCTCCATCGAGGAGCCCATG	1208	2290	TAGCCATTTCTGTCTCAACACCATCTCTAAGAAACAGCTTGAAGAGCTGTGTCTGGTC	2349
Qy	1210	GTGAAGGTGTTCAAGTCCGAGGCGCTGGCAGTGTGTGAGTGGGCTCGAGTCTC	1269	2289	TAGCCATTTCTGTCTCAACACCATCTCTAAGAAACAGCTTGAAGAGCTGTGTCTGGTC	2348
Db	1209	GTGAAGGTGTTCAAGTCCGAGGCGCTGGCAGTGTGTGAGTGGGCTCGAGTCTC	1268	2350	ATTGATTTTAACTAGAGGAGGAGGCTTAAACATGTACCCAGGAAACCATTTTAAACAGA	2409
Qy	1270	GGGGGCTTGGCTTACACAGGAGCTATCCGTACGAGGCGATCTCGTGAACACCCGCATC	1329	2349	ATTGATTTTAACTAGAGGAGGAGGCTTAAACATGTACCCAGGAAACCATTTTAAACAGA	2408
Db	1269	GGGGGCTTGGCTTACACAGGAGCTATCCGTACGAGGCGATCTCGTGAACACCCGCATC	1328	2410	ATATAAATGTCAAAATCTGTACTGTATAAAAAA 2449	
Qy	1330	CTCCTCATCTTCAGGGAACCAATGAGATTCTCCGATGTATACATCGCCCTGACGGGTG	1389	2409	ATATAAATGTCCCAATCAAAAAA 2448	
Db	1329	CTCCTCATCTTCAGGGAACCAATGAGATTCTCCGATGTATACATCGCCCTGACGGGTG	1388			
Qy	1390	CAGCATCGCGCGCATCTGACTACAGGAGTCCATGAGCTTAAACAGGCAAACTGAGC	1449			
Db	1389	CAGCATCGCGCGCATCTGACTACAGGAGTCCATGAGCTTAAACAGGCAAACTGAGC	1448			

RESULT 4
AF327351
LOCUS

AF327351 2411 bp mRNA linear PRI 30-DEC-2002

DEFINITION

Homo sapiens very-long-chain acyl-CoA dehydrogenase VLCAD mRNA, complete cds.

ACCESSION

AF327351

VERSION

1 GI:18028282

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 2411) Zhang, J., Zhang, W., Zou, D., Chen, G., Wan, T., Zhang, M. and Cao, X.

AUTHORS

Cloning and functional characterization of ACAD-9, a novel member

TITLE

of human acyl-CoA dehydrogenase family

JOURNAL

Biochem. Biophys. Res. Commun. 297 (4), 1033-1042 (2002)

MEDLINE

22246103

PUBMED

12359260

REFERENCE

2 (bases 1 to 2411)

AUTHORS

Zhang, W., Zhang, J., Li, N., Wan, T., Chen, T., Zhang, M. and Cao, X.

TITLE

Direct Submission

JOURNAL

Submitted (05-DEC-2000) Department of Immunology, Second Military

Medical University & Shanghai Brilliance Biotechnology Institute,

800 Xiangyin Rd., Shanghai 200433, P.R. China

FEATURES

source

1..2411
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
33..1898
/codon_start=1
/product="very-long-chain acyl-CoA dehydrogenase VLCAD"
/protein_id="AAL56011.1"
/db_xref="GI:18028283"
/translation="MSGGGLFRTTAAARACGLVSVSTANRLRLRTSPVRAFAKELF
LGKIKKEVPEPEVQSDELNINQFLGVEKFFTEEDVSRKIDQEGKIPDETLEKLE
SLGLQLQVPEYEGHIAFLCTEPASGDAASRATLSDEKKHILNGSVKWTNGG
QAKYLPKLSAGEHIAFLCTEPASGDAASRATLSDEKKHILNGSVKWTNGG
LANITVPKTVSDSDGSKDITAFIVERDFGVNKGPKEDKIGRSNTECVHFE
NTKIVENILGVEGDFKVMNLSNFRSVMGSLVAGLKLRIEMTAIYACTRQFNK
RUSEGLQERFALMAQKAYVMESYITAGMLDQPPDCSIEAAMVKVFSSEAOW
CVSEVLQIGGLGYTRPYERILRDRLITLIFETNEILRWYALTGLQHAGRLIT
RHELKQAKVSTYMDTVGRRLRSLGRTVDLGTNGHGVHPSLADSNKPEENTYCF
GRVETLLRFKFTIMEQLVKLVANILINLYGMTAVLSRPSIRIGLRNHDHEVL
LANTCEVAYLQNLFSLSQDKYAPENLDEIKKVSQILKEKRAYICAHPLDRTC"
BASE COUNT 612 a 604 c 669 g 526 t

ORIGIN

Query Match 97.5%; Score 2391.8; DB 9; Length 2411;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 2399; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 35 ACTGAGGCTGAGGCTGGGGAACATCGGCGACATGAGCGGCTGCGGGCTCTTCCTCGGCA 94

DB 1 ACTGAGGCTGAGGCTGGGGAACATCGGCGACATGAGCGGCTGCGGGCTCTTCCTCGGCA 60

QY 95 CCACGGCTGCGGCTGCTGCCCTGCGGGGTCTGGTCTCTACCGCAACCGCGGGCTAC 154

DB 61 CCACGGCTGCGGCTGCTGCCCTGCGGGGTCTGGTCTCTACCGCAACCGCGGGCTAC 120

QY 155 TCGCCACCGCCCGCTGCTAGAGCTTTCGCCCAAGAGCTTTTCTTAGGCAAAATCAAGA 214

DB 121 TCGCCACCGCCCGCTGCTAGAGCTTTCGCCCAAGAGCTTTTCTTAGGCAAAATCAAGA 180

QY 215 AGAAGAAGTTTTCCTAGAGCTTTCGCCCAAGAGCTTTTCTTAGGCAAAATCAAGA 274

DB 181 AGAAGAAGTTTTCCTAGAGCTTTCGCCCAAGAGCTTTTCTTAGGCAAAATCAAGA 240

QY 275 TCTTGGGACCGCTGGAAAAATCTTCTACTGAAGAGTGGGACTCCCGAAAAATGACACAG 334

DB 241 TCTTGGGACCGCTGGAAAAATCTTCTACTGAAGAGTGGGACTCCCGAAAAATGACACAG 300

QY 335 AAGGAAAAATCCAGATGAACCTTTCGAGAAATTTGAAGACCTTAGGGCTTTTGGGCTGC 394

DB 301 AAGGAAAAATCCAGATGAACCTTTCGAGAAATTTGAAGACCTTAGGGCTTTTGGGCTGC 360

395 AAGTCCCAAGAATAATGTCGCTGGCTTCTCCAAACACCATGTTACTCAAGACTAGGGG 454
|||||
361 AAGTCCCAAGAATAATGTCGCTGGCTTCTCCAAACACCATGTTACTCAAGACTAGGGG 420
|||||
455 AGATCATCAGCATGGATGGTCCATCACTGTGACCTGGCAGCGCACAGGCTATTGGCC 514
421 AGATCATCAGCATGGATGGTCCATCACTGTGACCTGGCAGCGCACAGGCTATTGGCC 480
|||||
515 TCAAGGGATCATCTTGGCTGGCACTGAGGACGAGAAAGCAAACTATTGCTTAACTGG 574
481 TCAAGGGATCATCTTGGCTGGCACTGAGGACGAGAAAGCAAACTATTGCTTAACTGG 540
|||||
575 CGTCCGGGAGCACATTTCAGCTTCTGCCCTCAGGAGCCAGCAGTGGGAGCCATGCAG. 634
|||||
541 CGTCCGGGAGCACATTTCAGCTTCTGCCCTCAGGAGCCAGCAGTGGGAGCCATGCAG 600
|||||
635 CCTCAATCCGAGGAGAGCCACACTAAGTGAAGACAAAGACACTACATCTCTAATGGCT 694
601 CCTCAATCCGAGGAGAGCCACACTAAGTGAAGACAAAGACACTACATCTCTAATGGCT 660
|||||
695 CCAAGGCTGCTGATTACTTAATGAGGACTGGCCCAATATTTTACTGTGTTGTCAAAGACTG 754
661 CCAAGGCTGCTGATTACTTAATGAGGACTGGCCCAATATTTTACTGTGTTGTCAAAGACTG 720
|||||
755 AGGTCGTTGATTCTGATGGATCAGTGAAGACAAACAAATCAGAGCAATTTAGGCAATTCGGGGCTCCCAACA 874
721 AGGTCGTTGATTCTGATGGATCAGTGAAGACAAACAAATCAGAGCAATTTAGGCAATTCGGGGCTCCCAACA 840
|||||
815 ACTTTGGTGGAGTCACTAATGGGAAACCCGAGCAATTTAGGCAATTCGGGGCTCCCAACA 874
781 ACTTTGGTGGAGTCACTAATGGGAAACCCGAGCAATTTAGGCAATTCGGGGCTCCCAACA 840
|||||
875 CTCTGAAGTCCATTTTGAACACACCAAGATACCTGTGGAACACATCTCTTGGAGAGTGC 934
841 CTCTGAAGTCCATTTTGAACACACCAAGATACCTGTGGAACACATCTCTTGGAGAGTGC 900
|||||
935 GAGATGGGTTAAGTGGCCATGAACATCTCAACAGCGCCGGTTTTCAGCATGGGAGCG 994
901 GAGATGGGTTAAGTGGCCATGAACATCTCAACAGCGCCGGTTTTCAGCATGGGAGCG 960
|||||
995 TCGTGGCTGGGCTGCTCAAGAGATTTGATTAAGATGACTGCTGAGTACGCTGCACAGGA 1054
961 TCGTGGCTGGGCTGCTCAAGAGATTTGATTAAGATGACTGCTGAGTACGCTGCACAGGA 1020
|||||
1055 AACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTTCAGGAGAAATTTGCACGTAGG 1114
1021 AACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTTCAGGAGAAATTTGCACGTAGG 1080
|||||
1115 CTGAGAGGCTTACGTCATGGAGATGACCTACCTACCTCAGCAGAGGATGCTGGAGCCAA 1174
1081 CTGAGAGGCTTACGTCATGGAGATGACCTACCTCAGCAGAGGATGCTGGAGCCAA 1140
|||||
1175 CTGGCTTTCCGACTCTCTCCTCAGGAGCCATGCTGAAGGTGTTTACGCTCCGAGGCGG 1234
1141 CTGGCTTTCCGACTCTCTCCTCAGGAGCCATGCTGAAGGTGTTTACGCTCCGAGGCGG 1200
|||||
1235 CTTGGCAGTGTGAGTGAAGCGCTGCAGATCCTCGGGGCTTGGGGCTTACACAGGGACT 1294
1201 CTTGGCAGTGTGAGTGAAGCGCTGCAGATCCTCGGGGCTTGGGGCTTACACAGGGACT 1260
|||||
1295 ATCCGTACGAGCGCATACTGCGTGACACCCGATCCTCTCATCTTCCAGGGAACCAATG 1354
1261 ATCCGTACGAGCGCATACTGCGTGACACCCGATCCTCTCATCTTCCAGGGAACCAATG 1320
|||||
1355 AGATTCTCGGATGTACATGCGCCCTGAGCGGTCTGACAGTATGCGGCGGCTCTCTGACTA 1414
1321 AGATTCTCGGATGTACATGCGCCCTGAGCGGTCTGACAGTATGCGGCGGCTCTCTGACTA 1380
|||||
1415 CCAGGATCCATGAGCTTAAACAGGCCAAAGTGAAGACAGTATGATGATACCGTTGGCGCGGA 1474
1381 CCAGGATCCATGAGCTTAAACAGGCCAAAGTGAAGACAGTATGATGATACCGTTGGCGCGGA 1440

1475 GGCCTGGGACCTCCCTGGGCGGAACTGTGGACCTGGGCTGACAGGCAACCATGGAGTTG 1534
|||||
1441 GGCCTGGGACCTCGCTGGGCGGAACTGTGGACCTGGGCTGACAGGCAACCATGGAGTTG 1500
|||||
1535 TGCACCCCACTCTTGGGACAGTGGCAACAACTTTGAGGAGAACCACTACTGCTTCGGCC 1594
|||||
1501 TGCACCCCACTCTTGGGACAGTGGCAACAACTTTGAGGAGAACCACTACTGCTTCGGCC 1560
|||||
1595 GGACCGTGGAGACACTGCTGCTCCGCTTTGGCAAGCAACATCATGAGGAGAGCTGGTAC 1654
|||||
1561 GGACCGTGGAGACACTGCTGCTCCGCTTTGGCAAGCAACATCATGAGGAGAGCTGGTAC 1620
|||||
1655 TGAAGCGGTGGCCACATCTCATCAACTGTATGGCATGACGCGCTGTCTCGCGG 1714
|||||
1621 TGAAGCGGTGGCCACATCTCATCAACTGTATGGCATGACGCGCTGTCTCGCGG 1680
|||||
1715 CCAGCGCTCCATCCGCACTTGGCTCCGCAACCAACGACGAGGTTCTTTGGCCAACA 1774
|||||
1681 CCAGCGCTCCATCCGCACTTGGCTCCGCAACCAACGACGAGGTTCTTTGGCCAACA 1740
|||||
1775 CCTTCTCGTGGAGCTTACTTGCAGATCTCTTCAAGCTCTCTCAGCTGGCAAGTATG 1834
|||||
1741 CCTTCTCGTGGAGCTTACTTGCAGATCTCTTCAAGCTCTCTCAGCTGGCAAGTATG 1800
|||||
1835 CTCAGAAAACCTAGATGACGAGATTAAAGAAAGTGTCCAGCAGATCCTTTGAGAACGAG 1894
|||||
1801 CTCAGAAAACCTAGATGACGAGATTAAAGAAAGTGTCCAGCAGATCCTTTGAGAACGAG 1860
|||||
1895 CCTATATCTGTGCCACCTCTTGACAGGACATGCTGAGCGAGGACAGTGTCCCTGC 1954
|||||
1861 CCTATATCTGTGCCACCTCTTGACAGGACATGCTGAGCGAGGACAGTGTCCCTGC 1920
|||||
1955 TACCGCCGCGCTTACCATGGCCGTTGCTGGATGACTTACTCTTTTTCAGAAGGT 2014
|||||
1921 TACCGCCGCGCTTACCATGGCCGTTGCTGGATGACTTACTCTTTTTCAGAAGGT 1980
|||||
2015 GTTGGGATATACAGGTTAAGCCTTTTGTTCCTGCTGACCTGAAGGTTTGTGCCT 2074
|||||
1981 GTTGGGATATACAGGTTAAGCCTTTTGTTCCTGCTGACCTGAAGGTTTGTGCCT 2040
|||||
2075 GGCCTGGAGAGCCTCTCCAGGTTTGTGCTGAGGACGTGCTCTTAACAGGACATC 2134
|||||
2041 GGCCTGGAGAGCCTCTCTCAGGTTTGTGCTGAGGACGTGCTCTTAACAGGACATC 2100
|||||
2135 ACAGCTTCTGAACCTGAGCGGAGAGAGAGATGAATGCTGACCCCTGAACTGGCGG 2194
|||||
2101 ACAGCTTCTGAACCTGAGCGGAGAGAGAGATGAATGCTGACCCCTGAACTGGCGG 2160
|||||
2195 TATCTGTGCTATGAGGAGACCATATGAGGAACTGGGGCTTATGCTGCTGCCCTCCAG 2254
|||||
2161 TATCTGTGCTATGAGGAGACCATATGAGGAACTGGGGCTTATGCTGCTGCCCTCCAG 2220
|||||
2255 GTGTGAGGTGGGTGGGACCTGTGTGAGGTGGATAGCCATTTCTGCTCAACACACAT 2314
|||||
2221 GTGTGAGGTGGGTGGGACCTGTGTGAGGTGGATAGCCATTTCTGCTCAACACACAT 2280
|||||
2315 TCTCTAAGAACAGCTTGAAGCTGTGCTGGTCAATTCATTTAACTAGAGAGAGG 2374
|||||
2281 TCTCTAAGAACAGCTTGAAGCTGTGCTGGTCAATTCATTTAACTAGAGAGAGG 2340
|||||
2375 ACTTAAACATGTACCCAGGACCAATTTAAACAAAGATATAAATGTCAATCTGTGTAC 2434
|||||
2341 ACTTAAACATGTACCCAGGACCAATTTAAACAAAGATATAAATGTCAATCTGTGTAC 2400
|||||
2435 TGTATAAAAAA 2445
|||||
2401 AAAAAA 2411
|||||

RESULT 5
BD158773 2399 bp DNA linear PAT 17-JAN-2003
LOCUS
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

BD158773
BD158773.1 GI:27864531
JP 2002191363-A/13616.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2399)
Ota,T., Isoqai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 13616 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/13616
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH key
Location/Qualifiers
FT CDS (47)..(1909).
1..2399
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 597 a 606 c 672 g 524 t
ORIGIN
Query Match 97.2%; Score 2382.2; DB 6; Length 2399;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2396; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
QY 21 GCTAAGAGGGGAGACTGAGGCTGGGCTGGGGAACATCGGGCAGCATGAGCGGCTGGCG 80
DB 1 GCTAAGAGGGGAGACTGAGGCTGGGCTGGGGAACATCGGGCAGCATGAGCGGCTGGCG 60
QY 81 GCTCTTCTTCGCGCACCGCTGCGGCTCGCTCGCGGGTCTGTGGTGTCTTACCGC 140
DB 61 GCTCTTCTTCGCGCACCGCTGCGGCTCGCTCGCGGGTCTGTGGTGTCTTACCGC 120
QY 141 GAACCGGGGCTACTGCGCACCGCTGCGGCTTTTCCCATTTCCAGAGTTAGCCAAAGACTTTTCT 200
DB 121 GAACCGGGGCTACTGCGCACCGCTGCGGCTTTTCCCATTTCCAGAGTTAGCCAAAGACTTTTCT 180
QY 201 AGCAAAATCAAGAAAGAAAGAGTTTTCCTTTCCCATTTCCAGAGTTAGCCAAAGACTTTTAA 260
DB 181 AGCAAAATCAAGAAAGAAAGAGTTTTCCTTTCCCATTTCCAGAGTTAGCCAAAGACTTTTAA 240
QY 261 TGAATCAATCACTTCTTGGGACCCCGTGGAAATTTCTTCACTGAGAGAGTGGACTCCG 320
DB 241 TGAATCAATCACTTCTTGGGACCCCGTGGAAATTTCTTCACTGAGAGAGTGGACTCCG 300
QY 321 AAAAATGACACGAGGAGAAATCCAGATGAACTTTGGGAAATTTGGAGAAATTTGAAGAGCTAG 380
DB 301 AAAAATGACACGAGGAGAAATCCAGATGAACTTTGGGAAATTTGAAGAGCTAG 360
QY 381 GCTTTTGGGCTCAAGTCCAGAAATATGTGGCTGGGCTTCTTCAACACCATGTA 440
DB 361 GCTTTTGGGCTCAAGTCCAGAAATATGTGGCTGGGCTTCTTCAACACCATGTA 420
QY 441 CTCAGACTAGGAGATCATCAGATGAGTGTGCTCATCTGTGACCTTGCAGCGCA 500
DB 421 CTCAGACTAGGAGATCATCAGATGAGTGTGCTCATCTGTGACCTTGCAGCGCA 480
QY 501 CCAGGCTATTGGCCTCAAGGGGATCATTTTGGCTGGCACTGAGGAGCAAGAACCAATA 560

Db 481 ||||| CCAGGCTATTGGCCCTCAAGGGGATCATCTTGGCTGGCACTGAGAGCAGAAAGCCAAATA 540
QY 561 CTTTGCTTAAACTGGCGTCCGGGAGCAGACATTGCAAGCCTTTCGCTCAGCGAGCCAGCCAG 620
Db 541 CTTTGCTTAAACTGGCGTCCGGGAGCAGACATTGCAAGCCTTTCGCTCAGCGAGCCAGCCAG 600
QY 621 TGGGAGCGATGAGCCTCAATCCGCGAGCAGAGCCACACTAAGTGAACACAGAGACACTA 680
Db 601 TGGGAGCGATGAGCCTCAATCCGCGAGCAGAGCCACACTAAGTGAACACAGAGACACTA 660
QY 681 CATCTCAATGGCTCCAAAGTCTCGATTACTTAATGGAGGACTGGCCCAATATTTTACTGT 740
Db 661 CATCTCAATGGCTCCAAAGTCTCGATTACTTAATGGAGGACTGGCCCAATATTTTACTGT 720
QY 741 GTTTGCAAGACTGAGGTGCTTGAATCTGATGGATCAGTGAAGAGACTGGCCCAATATTTTACTGT 800
Db 721 GTTTGCAAGACTGAGGTGCTTGAATCTGATGGATCAGTGAAGAGACTGGCCCAATATTTTACTGT 780
QY 801 CATAGTAGAAGAGACTTTGGTGAGTCACTAATGGGAAACCCGAAAGATAAATTAGGCAT 860
Db 781 CATAGTAGAAGAGACTTTGGTGAGTCACTAATGGGAAACCCGAAAGATAAATTAGGCAT 840
QY 861 TCGGGGCTCCAAACACTTGTCAAGTCCCAATTTTGAACACCAAGATACCTGTGGAAACAT 920
Db 841 TCGGGGCTCCAAACACTTGTCAAGTCCCAATTTTGAACACCAAGATACCTGTGGAAACAT 900
QY 921 CCTTGGAGAGGTCCGAGATGGGTTTAAAGTGGCCATGAACATCTCAACAGCGCCGGTT 980
Db 901 CCTTGGAGAGGTCCGAGATGGGTTTAAAGTGGCCATGAACATCTCAACAGCGCCGGTT 960
QY 981 CAGCATGGGAGCGTGGTGCGTGGCTGCTCAAGAGATTGATTGAATGACTGCTGAGTA 1040
Db 961 CAGCATGGGAGCGTGGTGCGTGGCTGCTCAAGAGATTGATTGAATGACTGCTGAGTA 1020
QY 1041 CGCCTGCACAAGGAAACAGTTTAAAGAGAGGTCACTGAATTTGGATTGATTTCAGAGAA 1100
Db 1021 CGCCTGCACAAGGAAACAGTTTAAAGAGAGGTCACTGAATTTGGATTGATTTCAGAGAA 1080
QY 1101 ATTTGACATGATGGCTCAGAAGCTTACGTCATGAGAGATGACTACCTCAGACAGG 1160
Db 1081 ATTTGACATGATGGCTCAGAAGCTTACGTCATGAGAGATGACTACCTCAGACAGG 1140
QY 1161 GATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGCCATGGTGAAGGTGT 1220
Db 1141 GATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGCCATGGTGAAGGTGT 1200
QY 1221 CAGCTCCGAGGCGGCTGGCAGTGTGAGTGAGCGCTGCAGATCCTCGGGGGCTTGGG 1280
Db 1201 CAGCTCCGAGGCGGCTGGCAGTGTGAGTGAGCGCTGCAGATCCTCGGGGGCTTGGG 1260
QY 1281 CTACACAAGGGACTATCCGTACGAGCGATCTGCGTGCACACCGCATCTCTCATCTT 1340
Db 1261 CTACACAAGGGACTATCCGTACGAGCGATCTGCGTGCACACCGCATCTCTCATCTT 1320
QY 1341 CGAGGGAACCAATGAGATTCTCCGGATGATACATCCGCTGACGGGTGACAGATCCCTCGGGGGCTTGGG 1400
Db 1321 CGAGGGAACCAATGAGATTCTCCGGATGATACATCCGCTGACGGGTGACAGATCCCTCGGGGGCTTGGG 1380
QY 1401 CGGCATCTGACTACAGGATCCAGTCTTAAACAGGCCAAAGTGAGCAGTCATGGA 1460
Db 1381 CGGCATCTGACTACAGGATCCAGTCTTAAACAGGCCAAAGTGAGCAGTCATGGA 1440
QY 1461 TACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAGCTGTGAGCCTGGGCTGACAGG 1520
Db 1441 TACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAGCTGTGAGCCTGGGCTGACAGG 1500
QY 1521 CAACCATGGAGTTGTGCACCCAGTCTTGGGAGAGTGCACCAAGTTTGGAGGAGAAC 1580
Db 1501 CAACCATGGAGTTGTGCACCCAGTCTTGGGAGAGTGCACCAAGTTTGGAGGAGAAC 1560
QY 1581 CTACTGCTTCCGCGGAGCGTGGAGACATGCTCCGCTTGGGAGACCATCATGGA 1640
|||||

Db 1561 CTACTGCTTCCGCGGAGCCGTGGAGACACTGCTGCTCCGCTTTGGCAAGACCATCATGGA 1620
QY 1641 GGAGCAGCTGGTACTGAAGCGGTGGCCAAACATCTCATCAACCTGTATGGCATGACGCG 1700
Db 1621 GGAGCAGCTGGTACTGAAGCGGTGGCCAAACATCTCATCAACCTGTATGGCATGACGCG 1680
QY 1701 CGTGTGTCGCGGGGAGCGGCTCCATCCGATTTGGGTCCGCAACACGACGACGAGGT 1760
Db 1681 CGTGTGTCGCGGGGAGCGGCTCCATCCGATTTGGGTCCGCAACACGACGACGAGGT 1740
QY 1761 TCTCTTGGCCACACACTTCTGGTGGAAAGCTTACTTTCAGAAATCTCTTCAGCCTCTCTCA 1820
Db 1741 TCTCTTGGCCACACACTTCTGGTGGAAAGCTTACTTTCAGAAATCTCTTCAGCCTCTCTCA 1800
QY 1821 GCTGGAAGATGATGCTCCAGAAACCTAGATGAGCAGATTAAAGAACTGTCACGACAGAT 1880
Db 1801 GCTGGAAGATGATGCTCCAGAAACCTAGATGAGCAGATTAAAGAACTGTCACGACAGAT 1860
QY 1881 CTTGAGAGGAGGCTATATCTGTGCCACCCCTCTGGACAGGACATGCTGAGGACGGG 1940
Db 1861 CTTGAGAGGAGGCTATATCTGTGCCACCCCTCTGGACAGGACATGCTGAGGACGGG 1920
QY 1941 ACAGTGTCCCTGCTTACCGCCGCGCTTACCCATGGCCGCTTGTGGATGACTGTACTC 2000
Db 1921 ACAGTGTCCCTGCTTACCGCCGCGCTTACCCATGGCCGCTTGTGGATGACTGTACTC 1980
QY 2001 TTTTTCAGAAAGGTGTTGGGATTATCAGAGTTAAGCCTTTTGTTCGCCGCTCTGCACCTG 2060
Db 1981 TTTTTCAGAAAGGTGTTGGGATTATCAGAGTTAAGCCTTTTGTTCGCCGCTCTGCACCTG 2040
QY 2061 AAGGTTGCTGCGCTGGCCTGGGAGAGCTTCTCCAGGTTTGGACCTGACGAGGAGTGTCT 2120
Db 2041 AAGGTTGCTGCGCTGGCCTGGGAGAGCTTCTCCAGGTTTGGACCTGACGAGGAGTGTCT 2100
QY 2121 CTAACAGGACCATCACAGCTTCTGAACCTGAGCGGAGAGAGAGAGAAATGCTGACCC 2180
Db 2101 CTAACAGGACCATCACAGCTTCTGAACCTGAGCG -- AGAGAGAGAAATGCTGACCC 2158
QY 2181 CTGGAACCTGGGGGTATTTCTGGTCAATTGAGGAGACACCATGATGGAACCTGGGCTTATG 2240
Db 2159 CTGGAACCTGGGGGTATTTCTGGTCAATTGAGGAGACACCATGATGGAACCTGGGCTTATG 2218
QY 2241 CTGCTGCTCCAGGCTGAGGTGGGTGGGAGCTGCTGTCAGGTGAGATGACCATTTCT 2300
Db 2219 CTGCTGCTCCAGGCTGAGGTGGGTGGGAGCTGCTGTCAGGTGAGATGACCATTTCT 2278
QY 2301 GCTCAACACACATTTCTTAAGAAACAGCTTGAAGAGCTCTGTCTGGGTGATTTAA 2360
Db 2279 GCTCAACACACATTTCTTAAGAAACAGCTTGAAGAGCTCTGTCTGGGTGATTTAA 2338
QY 2361 CTAGAAGCAGGACCTTAAACATGTACCAGGACCATTTAAACAAAGAAATATAATGT 2420
Db 2339 CTAGAAGCAGGACCTTAAACATGTACCAGGACCATTTAAACAAAGAAATATAATGT 2398
QY 2421 C 2421
Db 2399 C 2399

RESULT 6
LOUS AK024012
DEFINITION
Homo sapiens cDNA FLJ13950 fis, clone Y79AA1001048, weakly similar to ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC PRECURSOR (EC 1.3.99.-).
ACCESSION AK024012
VERSION AK024012.1 GI:10436257
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS	Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wadatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, N., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2399)
AUTHORS	Isogai, T. and Otsuki, T.
TITLE	Direct Submission
JOURNAL	Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES	Location/Qualifiers
source	1..2399
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="Y79AA1001048"
	/cell_line="Y79"
	/cell_type="retinoblastoma"
	/clone_lib="Y79AA1"
	/note="cloning vector: pME18SFL3"
	47..1912
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="BAB14775.1"
	/db_xref="GI:10436258"
CDS	translation="MSGGLFLRTAARACRGIVSTANRLLRTSPVRAFKEFL LGKIKKVFPEVPSQDELNEINQFLGPVEKPTFEBVDSKKIDQEGKIIDELEKIK SIGLFLQVPEYGGELGFSNTMYISRLGILISMDGSIITVLAHQIAIGKGLIAGTEG QKALYPLKLSAGEHIAFLPEPASGSDAISIRATLSDEDKHYILUNGSKVITNGG LANITVFARTEVDSQVSKDITAFIVERDFGVNGKPEDKLGIRGNTCEVHEE NFKIPVENLGEVGFQKVMNLSNFSRSMGVAGLRLRIEMTAETACRQFNK RLSEFLGKFAIMAKQAYVMSMTYLTAGMLDQPGFDCIEAAMKVFSSAAWQ CVSEALQILGGLYTRDYPERILDRILIFEGTNEILRMVIALTGLQHGRLIT RTHFKQAKVSTVMDTIGRLRDSIGRTVDLGTGNHGVHPSLADSANKFEENTYCF GRVTEILLRFGKIMERQLVLKRVANILINLXGMTAVLSRPSIRIGLRNHDHVEFL LANTCEVAYLQNLFLSLQDLKYPENLDEQIKVSOQILKRAYICAHPLDRFC"
BASE COUNT	597 a 606 c 672 g 524 t
ORIGIN	
Query Match	97.2%; Score 2382.2; DB 9; Length 2399;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2396; Conservative	0; Mismatches 3; Indels 2; Gaps 1;
QY	21 GCTAAGAGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGCATGAGCGGCTGGG 80
Db	
1	GCTAAGAGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGCATGAGCGGCTGGG 60
QY	81 GCTCTCTCCGCCACACGGCTGCGGCTCGCTCGCGGGGTCTGTTGCTCTACCGC 140
Db	
61	GCTCTCTCCGCCACACGGCTGCGGCTCGCTCGCGGGGTCTGTTGCTCTACCGC 120
QY	141 GAACCGCGGCTACTGCGCACCGCCGCTGTACGAGCTTTCGCCAAAGAGCTTTTCT 200
Db	
121	GAACCGCGGCTACTGCGCACCGCCGCTGTACGAGCTTTCGCCAAAGAGCTTTTCT 180
QY	201 AGGCAAAATCAGAGAAGAAGATTTTCCCATTTCCAGAGTTACCGAAGATGAACCTAA 260
Db	
181	AGGCAAAATCAGAGAAGAAGATTTTCCCATTTCCAGAGTTACCGAAGATGAACCTAA 240
QY	261 TGAATCAATCAGTTCTTGGGACCGCTGGAAAAATTTCTACTGAAGAGGTGGACTCCG 320
241	TGAAATCAATCAGTTCTTGGGACCGCTGGAAAAATTTCTACTGAAGAGGTGGACTCCG 300
QY	321 AAAAAATGACAGAGGAGGAAAAATCCAGATGAAACTTTTGGAGAAATTTGAAGACCTAGG 380
Db	
301	AAAAATGACAGAGGAGGAAAAATCCAGATGAAACTTTTGGAGAAATTTGAAGAGCTAGG 360
QY	381 GCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGGCTTGGCTTCTCCACACCATCTA 440
Db	
361	GCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGGCTTGGCTTCTCCACACCATCTA 420
QY	441 CTCAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCACTGTGACCCCTGGCAGCGCA 500
Db	
421	CTCAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCACTGTGACCCCTGGCAGCGCA 480
QY	501 CCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGCAAGAACCAATA 560
Db	
481	CCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGCAAGAACCAATA 540
QY	561 CTTCCTTAACCTGGGCTCGGGGAGCACATTCAGCCTTCTGCTCAGGAGCCAGCCAG 620
Db	
541	CTTCCTTAACCTGGGCTCGGGGAGCACATTCAGCCTTCTGCTCAGGAGCCAGCCAG 600
QY	621 TGGAGCGGATGAGCCTCAATCCGGAGCAGCCACACTAAGTGAAGCAAGAACACTA 680
Db	
601	TGGAGCGGATGAGCCTCAATCCGGAGCAGCCACACTAAGTGAAGCAAGAACACTA 660
QY	681 CATCTCAATGCTCCAAGGTCTGGATTACTAATGAGGACTGGCCCAATATTTTACTGT 740
Db	
661	CATCTCAATGCTCCAAGGTCTGGATTACTAATGAGGACTGGCCCAATATTTTACTGT 720
QY	741 GTTTGCAAGACTGAGGCTGTTGATTCTGATGGATCAGTGAAGCAAGAACCAATCAGCATT 800
Db	
721	GTTTGCAAGACTGAGGCTGTTGATTCTGATGGATCAGTGAAGCAAGAACCAATCAGCATT 780
QY	801 CATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAGATAAATTAGGCAT 860
Db	
781	CATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAGATAAATTAGGCAT 840
QY	861 TCGGGCTCCACACACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACACAT 920
Db	
841	TCGGGCTCCACACACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACACAT 900
QY	921 CCTTGGAGAGTGGAGATGGGTTTAAAGTGGCCATGAACATCTTCAACAGCGCGGTT 980
Db	
901	CCTTGGAGAGTGGAGATGGGTTTAAAGTGGCCATGAACATCTTCAACAGCGCGGTT 960
QY	981 CAGCTGGGAGGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGA 1040
Db	
961	CAGCTGGGAGGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGA 1020
QY	1041 CGCTGCAAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAA 1100
Db	
1021	CGCTGCAAGGAAACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAA 1080
QY	1101 ATTTGACTGATGGCTCAGAGGCTTACGTCATGGAGAGTATGACCTACCTCACAGCAG 1160
Db	
1081	ATTTGACTGATGGCTCAGAGGCTTACGTCATGGAGAGTATGACCTACCTCACAGCAG 1140
QY	1161 GATCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGCCATGGTGAAGTGT 1220
Db	
1141	GATCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGCCATGGTGAAGTGT 1200
QY	1221 CAGCTCGAGCGCGCTGGCAGTGTGTAGTGAAGCGCTGCAGATCCTCGGGGCTTGG 1280
Db	
1201	CAGCTCGAGCGCGCTGGCAGTGTGTAGTGAAGCGCTGCAGATCCTCGGGGCTTGG 1260
QY	1281 CTACACAAGGAGCTATCCGTACGAGGCGCATACTCGGTGACACCCGCATCCTCTCATCT 1340
Db	
1261	CTACACAAGGAGCTATCCGTACGAGGCGCATATTGCGGTGACACCCGCATCCTCTCATCT 1320
QY	1341 CGAGGGAACCAATGAGATTCTCGGATGTACATCGCCCTGACGGGTCTCGAGCATCCCG 1400
Db	
1321	CGAGGGAACCAATGAGATTCTCGGATGTACATCGCCCTGACGGGTCTCGAGCATCCCG 1380

ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
AUTHORS	1 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Sasaki, N. NEDO human cDNA sequencing project	
TITLE	Unpublished	
JOURNAL	2 (bases 1 to 2400)	
AUTHORS	Isogai, T. and Otsuki, T.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.	
FEATURES	Location/Qualifiers	
source	1..2400	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="NT2RM2001700"	
	/cell_line="NT2"	
	/cell_type="teratocarcinoma"	
	/clone_lib="NT2RM2"	
	/note="cloning vector: pME18SFL3-mRNA from uninduced NT2 neuronal precursor cells."	
CDS	498..1994	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="BAB14104.1"	
	/db_xref="GI:10434033"	
	/translation="MYSRIEIIISMDGSITVLAHQAIIGLKILLAGTERQKAKYLP KLASGEHAACFLTPASGDAASIRSLSEDKKHVILNGSKVWITNGLANFTV FAKTEVDSQSVKDKITAFIVERDFGVNGKPKDLIRGNTCEVHPKIPVE NILGEVGGDFKVMNLNSRFSMGSVVGALIKRLIEMTAIEYACTRQFNKRLEFGL IOEKPALMAQKAYVNESMTYLTAGMLDQPGPDCSIEAMVKVFSSEAQCVCSEALQ ILGGIYTRDYPYRILPDRILLIFEGTNEILRMVIALTGLOHAGRLITRIHELQ AKVSTVMDTVGRRLLRDSLGRTVLDLGNHGVHPSLADSKANFEENTYCFGRVETL LLRFKGTIMESQVLKRVANILINLYGMTAVLSRSIRIGLRNHDHEVLLANTFCV EAYLQNLFSQLDKYAPENLDEQIKKVSQOILEKRAYICAHPLDRTIC"	
BASE COUNT	570 a 629 c 685 g 516 t	
ORIGIN	<p>Query Match 86.3%; Score 2115.4; DB 9; Length 2400; Best Local Similarity 94.2%; Pred. No. 0; Matches 2262; Conservative 0; Mismatches 56; Indels 83; Gaps 3;</p> <p>QY 21 GCTAAGAGGGAGACTGAGGCTGGGGGAACATCGGGAGCATCGGCGCTCGCG 80 DB 1 GCTAAGAGGGAGACTGAGGCTGGGGGAACATCGGGAGCATCGGCGCTCGCG 60 </p> <p>QY 81 GCTCTTCTCGGCACACGCGTGGCGTCTGCTCGCGGGGTCTGGTCTCTACCGC 140 DB 61 GCTCTTCTCGGCACACGCGTGGCGTCTGCTCGCGGGGTCTGGTCTCTACCGC 120 </p> <p>QY 141 GAACCGCGGCTACTGCGGCACACGCGCTGTACGAGCTTCGCCAAGAGCTTTTCT 200 DB 121 GAACCGCGGCTACTGCGGCACACGCGCTGTACGAGCTTCGCCAAGAGCTTTTCT 180 </p> <p>QY 201 AGGCAAAATCAAGAGAAA----- DB 181 AGGCAAAATCAAGAGAAA----- </p> <p>QY 220 -----GAAGTTTCCCATTTCCAGAGATTTAGCCAAAGAT 252 DB -----GAAGTTTCCCATTTCCAGAGATTTAGCCAAAGAT 230 </p>	

Db	241	TTTACCCTCAGCTGCAAGACTGGTGTGAACCTTTGTGAGATTTCCCAACCTGCCAGAGA	300
QY	253	GAACCT-----TAATGAATCAATCACTTTTGGGACCCCGTGAATAATTTCT	298
Db	301	GATACACCTCGCGCGAGGGCTGTAACTCCCGATTTCCTGAGTTCCAGGAACCTT	360
QY	299	TCACCTGAAGAGGTGGACTCCCCAAAATTTGACACAGAGGAGGAAATCCACATGAACTT	358
Db	361	CCACAGAGAAAGGTGACTCCCGAAAATTTGACAGGAGGAGGAAATCCAGATGAACTT	420
QY	359	TGGAGAAATTTGAAGACCTTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGCC	418
Db	421	TGGAGAAATTTGAAGACCTTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGCC	480
QY	419	TGGGCTTCTTCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCA	478
Db	481	TGGGCTTCTTCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCA	540
QY	479	TCACTGTACCTCGCAGCGCACAGGCTATTGGCCCTCAAGGGGATCATCTTGGCTGCCA	538
Db	541	TCACTGTACCTCGCAGCGCACAGGCTATTGGCCCTCAAGGGGATCATCTTGGCTGCCA	600
QY	539	CTGAGGAGCAAAAGCCAAATACTTGCCTAACTGGCGTCCGGGGAGCACATTTGCAGCCT	598
Db	601	CTGAGGAGCAAAAGCCAAATACTTGCCTAACTGGCGTCCGGGGAGCACATTTGCAGCCT	660
QY	599	TCTGCTCAGCGGACCCAGCGAGTGGAGCGATGAGCCTCAATCCCGAGCAGACCCACAC	658
Db	661	TCTGCTCAGCGGACCCAGCGAGTGGAGCGATGAGCCTCAATCCCGAGCAGACCCACAC	720
QY	659	TAAGTGAAGACAAGAACACTACATCTCAATGGCTCCAAAGTCTCGATTAATGGAG	718
Db	721	TAAGTGAAGACAAGAACACTACATCTCAATGGCTCCAAAGTCTCGATTAATGGAG	780
QY	719	GACTGGCCAAATTTTACTGTGTTCGAAAGACTGAGGTCGTTGATCTGATGGATCAG	778
Db	781	GACTGGCCAAATTTTACTGTGTTCGAAAGACTGAGGTCGTTGATCTGATGGATCAG	840
QY	779	TGAAGACAAATACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGA	838
Db	841	TGAAGACAAATACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGA	900
QY	839	AACCCGAAGATAAATTTAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTGAAACA	898
Db	901	AACCCGAAGATAAATTTAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTGAAACA	960
QY	899	CAAAGATACCTGTGGAAAACATCTTTGGAGAGGTGGGATGGGTTTAAAGTGGCCATGA	958
Db	961	CAAAGATACCTGTGGAAAACATCTTTGGAGAGGTGGGATGGGTTTAAAGTGGCCATGA	1020
QY	959	ACATCTCAACAGCGCGGTTTCAGCATGGGCGAGCGTGTGGCTGGGCTGCTCAAGAGAT	1018
Db	1021	ACATCTCAACAGCGCGGTTTCAGCATGGGCGAGCGTGTGGCTGGGCTGCTCAAGAGAT	1080
QY	1019	TGATTGAATGACTGCTGAGTACGCTGCACAAAGGAAACAGTTTAAAGAGAGGCTCAGTG	1078
Db	1081	TGATTGAATGACTGCTGAGTACGCTGCACAAAGGAAACAGTTTAAAGAGAGGCTCAGTG	1140
QY	1079	AATTTGGATTGATTGAGAGAAATTTGCACATGGCTCAGAGGCTTACGTCATGGAGA	1138
Db	1141	AATTTGGATTGATTGAGAGAAATTTGCACATGGCTCAGAGGCTTACGTCATGGAGA	1200
QY	1139	GTATGACCTTACCTACAGCAGGATGTGGACCAACCTGGCTTTCCCGACTGCTCCATCG	1198
Db	1201	GTATGACCTTACCTACAGCAGGATGTGGACCAACCTGGCTTTCCCGACTGCTCCATCG	1260
QY	1199	AGGAGCCATGGTGAAGGTGTTGAGTCCGAGCGGCTGGCAGTGTGTGAGTGAGGCGC	1258
Db	1261	AGGAGCCATGGTGAAGGTGTTGAGTCCGAGCGGCTGGCAGTGTGTGAGTGAGGCGC	1320
QY	1259	TGCAGATCCTCGGGGGCTTGGCTTACACAAAGGACTATCCGTAGCAGCGCATCTGCGTG	1318
Db	1321	TGCAGATCCTCGGGGGCTTGGCTTACACAAAGGACTATCCGTAGCAGCGCATCTGCGTG	1380

QY	1319	ACACCGCATCCTCCTCATCTTCGAGGGAACCAATGAGATTCTCGGATGTACATCGGCC	1378
Db	1381	ACACCGCATCCTCCTCATCTTCGAGGGAACCAATGAGATTCTCGGATGTACATCGGCC	1440
QY	1379	TGACGGGTCTGAGCATGCCGCCCATCTCTGACTACAGGATCCATGAGTTAAACAGG	1438
Db	1441	TGACGGGTCTGAGCATGCCGCCCATCTCTGACTACAGGATCCATGAGTTAAACAGG	1500
QY	1439	CCAAAGTGAGCAGTATGATACCGTTGGCGGAGGCTTGGGACTCCCTGGGCCGAA	1498
Db	1501	CCAAAGTGAGCAGTATGATACCGTTGGCGGAGGCTTGGGACTCCCTGGGCCGAA	1560
QY	1499	CTGTGACCTGGGGCTGACAGGCAACCATGAGTTGTGCACCCAGTCTTGGGACAGT	1558
Db	1561	CTGTGACCTGGGGCTGACAGGCAACCATGAGTTGTGCACCCAGTCTTGGGACAGT	1620
QY	1559	CCAAAGTTGAGGAGAACACTACTGCTTCGGCGGACCGTGGAGACACTGTCGTCC	1618
Db	1621	CCAAAGTTGAGGAGAACACTACTGCTTCGGCGGACCGTGGAGACACTGTCGTCC	1680
QY	1619	GCTTTGGCAAGACCATCATGGAGGAGCCTGCTGACTGAAGGGGTGGCCACATCTCA	1678
Db	1681	GCTTTGGCAAGACCATCATGGAGGAGCCTGCTGACTGAAGGGGTGGCCACATCTCA	1740
QY	1679	TCAACCTGTATGATGATGAGCGGCTGCTGCGGGCCAGCGCTCCATCCGCAITGGC	1738
Db	1741	TCAACCTGTATGATGATGAGCGGCTGCTGCGGGCCAGCGCTCCATCCGCAITGGC	1800
QY	1739	TCCGCAACCGACCGAGGCTTCTCTTGGCCAAACCTTCTGCGTGGAGCTTACTTGC	1798
Db	1801	TCCGCAACCGACCGAGGCTTCTCTTGGCCAAACCTTCTGCGTGGAGCTTACTTGC	1860
QY	1799	AGATATCTTCAAGCTCTCTCAGTGGACAGTATGCTCCAGAAAACCTAGATGACAGA	1858
Db	1861	AGATATCTTCAAGCTCTCTCAGTGGACAGTATGCTCCAGAAAACCTAGATGACAGA	1920
QY	1859	TTAAGAAAGTGTCCAGCAGATCCTTGAAGGAGGAGCTATATCTGTGCCACCTCTGG	1918
Db	1921	TTAAGAAAGTGTCCAGCAGATCCTTGAAGGAGGAGCTATATCTGTGCCACCTCTGG	1980
QY	1919	ACAGGATGTGAGGAGGGGACAGTGTCCCTGTACCGCCGCCCTTACCATGGCC	1978
Db	1981	ACAGGATGTGAGGAGGGGACAGTGTCCCTGTACCGCCGCCCTTACCATGGCC	2040
QY	1979	CGTTGCTGATGACTGTACTCTTTTTCAGAAAGTGTGGGATTAACAGGTTAAGCC	2038
Db	2041	CGTTGCTGATGACTGTACTCTTTTTCAGAAAGTGTGGGATTAACAGGTTAAGCC	2100
QY	2039	TTTTGTTCCCGTGTGCACCTGAAGGTTGTGCGCTGGCTGGGAGAGCTCTTCCAGT	2098
Db	2101	TTTTGTTCCCGTGTGCACCTGAAGGTTGTGCGCTGGCTGGGAGAGCTCTTCCAGT	2160
QY	2099	TTTGACCTGCGAGGAGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCGGAGA	2158
Db	2161	TTTGACCTGCGAGGAGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCGGAGA	2220
QY	2159	GAGAAATGGAATGCTGACCCCTGGAACCTGGCGGTATTTGTCATTTGAGGAGACAC	2218
Db	2221	GAGAAATGGAATGCTGACCCCTGGAACCTGGCGGTATTTGTCATTTGAGGAGACAC	2279
QY	2219	ATAGTGGAACTGGGGCTTATGCTGCTCCAGGTTGTGAGTGGGTGGGACCTGTG	2278
Db	2280	ATAGTGGAACTGGGGCTTATGCTGCTCCAGGTTGTGAGTGGGTGGGACCTGTG	2339
QY	2279	TCAGTGTGATAGGCAATTTCTGCTCAACACACATCTCTAAGAAACAGCTTGAAGCT	2338
Db	2340	TCAGTGTGATAGGCAATTTCTGCTCAACACACATCTCTAAGAAACAGCTTGAAGCT	2399
QY	2339	C 2339	
Db	2400	C 2400	

RESULT 9

AX394749

LOCUS

DEFINITION

AX394749

ACCESSION

AX394749.1

VERSION

GI:21065828

KEYWORDS

Source

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..1863

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

1..>1863

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD32050.1"

/db_xref="GI:21065828"

/translation="MSGGGLPLRTAAARACRGLVSTANRRLLRTPPVRAFAKELF

LGIKKKEVFPFPEVSDLEINQFPGVKEKFTTEVDSEKIDQEGKIPELTLEKLK

SLGLFGLQVPEYGGGLSNTMYSLGELIISMDGSIITVTLAAHQAIGLIGLAGTGE

QKAKYLPKLASGEHIAAFCTEPASGDAASIRSRATLSDEKHLIINGSKVITNGE

LANLTFVFAKTVSDGSKVDKITAFIVDERDVGVTNGKPEDLGRINSGNTCEVFE

NTKTPVENILGEVGDGPKAVNINLSRFSMSVAGLLKRLIEMAYKVFSSAAWQ

RLSEFGLIQEFALMAKAVYVESWTVLTAGMLDQPGPPDCSIEAAMVKVFSSEAQQ

CVSEALQILGGLGYTRDIPYERILDRILLIFEGTNEILRMVIALSGLOHARILIT

RIHELKQAKVSTVMDTSGRLRDSLRITVDLGLIGNHGVHPSLADSNKFEENTYCF

GRVTETLLREFKTIMEQLVKRVALINLYGMTAVLSRASIRIGLRNHDHEVL

LANTPCVEAYLQNLFLSLQDLKYPENIDQIKVSOQILEKRAYICAHPLDRTC"

BASE COUNT

471 a 481 c 519 g 392 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 ATGAGCGGCTCGGGGCTCTTCTGCGCACCGCTGCGGCTGCTGCGGCGGTCTG 126

Db 1 ATGAGCGGCTCGGGGCTCTTCTGCGCACCGCTGCGGCTGCTGCGGCGGTCTG 60

QY 127 GTGGTCTCTACCGGAAACCGCGGCTGCTGCGCACCGCTGCTGCGGCGGTCTG 186

Db 61 GTGGTCTCTACCGGAAACCGCGGCTGCTGCGCACCGCTGCTGCGGCGGTCTG 120

QY 187 AAAGAGCTTTCTCTAGGCAAAATCAAGAAGAAGAGTTTCCCATTTCCAGAAAGTTAGC 246

Db 121 AAAGAGCTTTCTCTAGGCAAAATCAAGAAGAAGAGTTTCCCATTTCCAGAAAGTTAGC 180

QY 247 CAAGATGAACCTTAATGAAATCAATCAGTTCTTGGGACCGCTGGAATAATCTTCACTGAA 306

Db 181 CAAGATGAACCTTAATGAAATCAATCAGTTCTTGGGACCGCTGGAATAATCTTCACTGAA 240

QY 307 GAGGTGGACTCCCGAAATAATTGACAGGAGGAGAAATCCAGATGAAACTTTGGAGAAA 366

Db 241 GAGGTGGACTCCCGAAATAATTGACAGGAGGAGAAATCCAGATGAAACTTTGGAGAAA 300

QY 367 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCGAGAAATATATGGTGGCGTGGCGCTTC 426

Db 301 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCGAGAAATATATGGTGGCGTGGCGCTTC 360

QY 427 TCCAACACCATGTACTCAAGCTAGGGGAGATCATCAGCATGATGGTCCCATCTGTC 486

Db 361 TCCAACACCATGTACTCAAGCTAGGGGAGATCATCAGCATGATGGTCCCATCTGTC 420

QY 487 ACCCTGGCAGGCGACAGGCTATTGGCCTCAAGGGATCATCTTGGCTGGCACTGAGGAG 546
 Db 421 ACCCTGGCAGGCGACAGGCTATTGGCCTCAAGGGATCATCTTGGCTGGCACTGAGGAG 480
 QY 547 CAGAAAGCCAAATACTTGCCTAAACTGGCGTCCGGGAGGACACATTTGCGAGCTTCTGCCTC 606
 Db 481 CAGAAAGCCAAATACTTGCCTAAACTGGCGTCCGGGAGGACACATTTGCGAGCTTCTGCCTC 540
 QY 607 ACGGAGCCAGCAGTGGGAGGATGACAGCCTCAATCCGGAGGAGGACACATTTGCGAGCTTCTGCCTC 666
 Db 541 ACGGAGCCAGCAGTGGGAGGATGACAGCCTCAATCCGGAGGAGGACACATTTGCGAGCTTCTGCCTC 600
 QY 667 GACAAGAGCACTACATCTCTCAATGGCTCCAAGGCTCTGGATTACTAATGAGGAGCTGGCC 726
 Db 601 GACAAGAGCACTACATCTCTCAATGGCTCCAAGGCTCTGGATTACTAATGAGGAGCTGGCC 660
 QY 727 AATATTTTACTGTGTTTGCAGAGCTGAGGCTGGTGGATTCTGATGATCTGATGATCAAGTGA 786
 Db 661 AATATTTTACTGTGTTTGCAGAGCTGAGGCTGGTGGATTCTGATGATCTGATGATCAAGTGA 720
 QY 787 AAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGCTCACTAATTTGAAAACACCAAGATA 906
 Db 721 AAAATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGCTCACTAATTTGAAAACACCAAGATA 846
 QY 847 GATAAATTAGGCATTCGGGGCTCCAACTGTTGAAGTCCATTTTGAAGAACCAAGATA 840
 Db 781 GATAAATTAGGCATTCGGGGCTCCAACTGTTGAAGTCCATTTTGAAGAACCAAGATA 800
 QY 907 CCTGTGGAAGCAATCCTTTGGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGACATCTCTC 966
 Db 841 CCTGTGGAAGCAATCCTTTGGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGACATCTCTC 900
 QY 967 AACAGCGCGCGTTCAGCATGGGAGCGCTGGTGGCTGCTCAAGAGATTGATTGAA 1026
 Db 901 AACAGCGCGCGTTCAGCATGGGAGCGCTGGTGGCTGCTCAAGAGATTGATTGAA 960
 QY 1027 ATGACTGCTCAGTACCCCTGACAGAGAAAGATTAAAGAGGCTCAGTGAATTTGGA 1086
 Db 961 ATGACTGCTCAGTACCCCTGACAGAGAAAGATTAAAGAGGCTCAGTGAATTTGGA 1020
 QY 1087 TTGATTTCAGGAGAAATTTGCATCTGATGGCTCAGAGGCTTACGTATGAGAGTATGACC 1146
 Db 1021 TTGATTTCAGGAGAAATTTGCATCTGATGGCTCAGAGGCTTACGTATGAGAGTATGACC 1080
 QY 1147 TACCTCACAGCAGGAGTGTGACCAACCTGCTTCCCGACTGCTCCATCGAGGAGCC 1206
 Db 1081 TACCTCACAGCAGGAGTGTGACCAACCTGCTTCCCGACTGCTCCATCGAGGAGCC 1140
 QY 1207 ATGGTGAAGGTGTTGAGTCCGAGCGCGCTGGCAGTGTGAGTGAGGCGCTGCAGATC 1266
 Db 1141 ATGGTGAAGGTGTTGAGTCCGAGCGCGCTGGCAGTGTGAGTGAGGCGCTGCAGATC 1200
 QY 1267 CTCGGGGGTTGGCTTACACAGGAGCTATCCGTACGAGCGCATCTCGTGCACACCGC 1326
 Db 1201 CTCGGGGGTTGGCTTACACAGGAGCTATCCGTACGAGCGCATCTCGTGCACACCGC 1260
 QY 1327 ATCTCTCTCATCTTCAGGAGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGT 1386
 Db 1261 ATCTCTCTCATCTTCAGGAGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGT 1320
 QY 1387 CTCGAGCATCGCGGCGCATCTGACTACCAAGGATCCATGAGTCTAAACAGGCCAAAGTG 1446
 Db 1321 CTCGAGCATCGCGGCGCATCTGACTACCAAGGATCCATGAGTCTAAACAGGCCAAAGTG 1380
 QY 1447 AGCACATCATGATACCGTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1506
 Db 1381 AGCACATCATGATACCGTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1440
 QY 1507 CTGGGGCTGACAGGCAACCATGGAGTTGTGACCCAGCTCTTGGGAGCAGTGCCAAAG 1566
 Db 1441 CTGGGGCTGACAGGCAACCATGGAGTTGTGACCCAGCTCTTGGGAGCAGTGCCAAAG 1500
 QY 1567 TTTGAGGAGAAACCTACTGCTTCGGCGGACCGCTGGAGACACTGCTGCTCGCTTTGGC 1626

Db 1501 TTTGAGGAGAAACACCTACTGCTTTCGGCGGACCGTGGAGACACTGCTGCTCGCTTTGGC 1560
 QY 1627 AAGACCATCATGAGGAGGAGCTGGTACTGAAGCGGGTGGCCAAACATCTCTCATCAACCTG 1686
 Db 1561 AAGACCATCATGAGGAGGAGCTGGTACTGAAGCGGGTGGCCAAACATCTCTCATCAACCTG 1620
 QY 1687 TATGGCATGACGCGCGCTGCTGCGGGCCAGCGCTCCATCCGCAATTTGGCTCGGCAAC 1746
 Db 1621 TATGGCATGACGCGCGCTGCTGCGGGCCAGCGCTCCATCCGCAATTTGGCTCGGCAAC 1680
 QY 1747 CAGGAGCAGGAGGTTCTCTTGGCCAAACACCTTCTGCGTGGAGCTTACTTGCAGAACTC 1806
 Db 1681 CAGGAGCAGGAGGTTCTCTTGGCCAAACACCTTCTGCGTGGAGCTTACTTGCAGAACTC 1740
 QY 1807 TTAGCCCTCTCTCAGCTGAGCAAGTATGCTCCAGAAACCTTCCAGAAACCTTAAAGAA 1866
 Db 1741 TTAGCCCTCTCTCAGCTGAGCAAGTATGCTCCAGAAACCTTCCAGAAACCTTAAAGAA 1800
 QY 1867 GTGTCCCAGCAGATCTTTGAGAAGCGAGCTTATATCTGTGCCACCCCTCTTGACAGGACA 1926
 Db 1801 GTGTCCCAGCAGATCTTTGAGAAGCGAGCTTATATCTGTGCCACCCCTCTTGACAGGACA 1860
 QY 1927 TGC 1929
 Db 1861 TGC 1863

BC001817 1494 bp mRNA linear PRI 12-JUL-2001
 Homo sapiens, Similar to NP0002 protein, clone MGC:3368
 IMAGE:3530094, mRNA, complete cds.
 BC001817
 MGC.
 BC001817.1 GI:12804754
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Strausberg, R.
 Direct Submission
 Submitted (29-JAN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 Contact: amadanesystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://lmgc.llnl.gov>
 Series: IRAL Plate: 11 Row: k Column: 19.
 Location/Qualifiers
 1. 1494
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:3368 IMAGE:3530094"
 /tissue_type="Muscle, rhabdomyosarcoma"
 /clone_lib="NIH MGC_17"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 475. 1041

FEATURES
 source
 1. 1494

CDS

/codon_start=1	
/product="Similar to NP0002 protein"	
/protein_id="AA01817.1"	
/db_xref="GI:1280475"	
/translation="MYALTLGQHGRIITRIHELKQAKVSTVMDTVGRRLRSLGR	
TVDLGLTGNHGVHPVSLADSAKPFENTYCFGRITVETLILRFKGTIMEQVLKRVAN	
ILINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAVLQNLFSQLDKVAPEN	
LDEQIKVSSQILERAYICAPLDRIC"	
BASE COUNT	370 a 386 c 422 g 316 t
ORIGIN	
Query Match 58.7%; Score 1439; DB 9; Length 1494;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 1439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	901 AAGATACCTGTGAAACATCCTTTGAGAGGTCGGAGATGGGTTTAAAGTGCCCATGAAC 960
Db	10 AAGATACCTGTGAAACATCCTTTGAGAGGTCGGAGATGGGTTTAAAGTGCCCATGAAC 69
Qy	961 ATCTCAACAGCGCGGTTTACGATGGCGAGGTCGTGGCTGGGCTGCTCAAGAGATTG 1020
Db	70 ATCTCAACAGCGCGGTTTACGATGGCGAGGTCGTGGCTGGGCTGCTCAAGAGATTG 129
Qy	1021 ATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAA 1080
Db	130 ATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAA 189
Qy	1081 TTTGGATTGATTCAGGAGAAATTTGCACATGCTGCTCAGAGGCTTACGTCATGGAGAGT 1140
Db	190 TTTGGATTGATTCAGGAGAAATTTGCACATGCTGCTCAGAGGCTTACGTCATGGAGAGT 249
Qy	1141 ATGACCTACCTCAGCAGGAGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAG 1200
Db	250 ATGACCTACCTCAGCAGGAGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAG 309
Qy	1201 GCAGCATGTTGAAGTGTTCAGTCCGAGCGCCCTGGCAGTGTGTGAGTGAGGCGCTG 1260
Db	310 GCAGCATGTTGAAGTGTTCAGTCCGAGCGCCCTGGCAGTGTGTGAGTGAGGCGCTG 369
Qy	1261 CAGATCTCTCGGGGCTTGGGTACACAAAGGACTATCCGTACGAGCGCATACACTGCGTGAC 1320
Db	370 CAGATCTCTCGGGGCTTGGGTACACAAAGGACTATCCGTACGAGCGCATACACTGCGTGAC 429
Qy	1321 ACCCGCATCTCTCTATCTTCGAGGAACCAATGAGATTCTCCGGATGTACATCGCCCTG 1380
Db	430 ACCCGCATCTCTCTATCTTCGAGGAACCAATGAGATTCTCCGGATGTACATCGCCCTG 489
Qy	1381 ACGGTCGTGACGATCGCGCGCATCTGACTACGAGATCCATGAGCTTAAACAGGCC 1440
Db	490 ACGGTCGTGACGATCGCGCGCATCTGACTACGAGATCCATGAGCTTAAACAGGCC 549
Qy	1441 AAGTGAGCACAGTCATGATACCGTTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAAC 1500
Db	550 AAGTGAGCACAGTCATGATACCGTTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAAC 609
Qy	1501 GTGGACCTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGCTCTTGGCGAGTGCC 1560
Db	610 GTGGACCTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGCTCTTGGCGAGTGCC 669
Qy	1561 AACAAGTTTGAGGAGAACACCTACTGCTTTCGGCGGACCGGTGGAGACACTGCTGCTCGC 1620
Db	670 AACAAGTTTGAGGAGAACACCTACTGCTTTCGGCGGACCGGTGGAGACACTGCTGCTCGC 729
Qy	1621 TTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACATCTCATC 1680
Db	730 TTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACATCTCATC 789
Qy	1681 AACCTGTATGGCATGACGCGCTGTGTGCGGGCGAGCGGCTCCATCGCATTTGGGCTC 1740
Db	790 AACCTGTATGGCATGACGCGCTGTGTGCGGGCGAGCGGCTCCATCGCATTTGGGCTC 849
Qy	1741 CGCAACACGACACAGAGTCTCTTTGGCCAAACACCTTCTGCGTGGAGGCTTACTTGAC 1800

Db	850 CGCAACACGACGACGAGGTTCTTCTGGCCAAACACCTTCTGCGTGGAGCTTACTTGAC 909
Qy	1801 AATCTCTTACGCTCTCTCAGCTGGACAAGTATGCTCCAGAAAACCTAGATGACAGATT 1860
Db	910 AATCTCTTACGCTCTCTCAGCTGGACAAGTATGCTCCAGAAAACCTAGATGACAGATT 969
Qy	1861 AAGAAAGTGTCCCAAGCAGATCTTTGAGAAAGGAGCCTATATCTGTGCCACCTCTTGAC 1920
Db	970 AAGAAAGTGTCCCAAGCAGATCTTTGAGAAAGGAGCCTATATCTGTGCCACCTCTTGAC 1029
Qy	1921 AGGACATGCTGAGGAGGAGACAGTCTCCCTGCTAGCGCGCGCCCTACCCATGGCCCG 1980
Db	1030 AGGACATGCTGAGGAGGAGACAGTCTCCCTGCTAGCGCGCGCCCTACCCATGGCCCG 1089
Qy	1981 TTGCTGGATGACTGTTACTCTTTTTCAGAAAGGTTGTTGGGATTTATCAGAGTTAAGCCTT 2040
Db	1090 TTGCTGGATGACTGTTACTCTTTTTCAGAAAGGTTGTTGGGATTTATCAGAGTTAAGCCTT 1149
Qy	2041 TTGTTCCCGCTGTCACCTGAAGGTTGTCGCCCTGGGCTGGGAGAGCTCTTCCAGGTTT 2100
Db	1150 TTGTTCCCGCTGTCACCTGAAGGTTGTCGCCCTGGGCTGGGAGAGCTCTTCCAGGTTT 1209
Qy	2101 TGACCTGTCAGGAGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCGGAGAGA 2160
Db	1210 TGACCTGTCAGGAGTCTCTTAACAGGACCATCACAGCTTCTGAACCTGAGCGGAGAGA 1269
Qy	2161 GAGATGGAATTTGCTGACCCCTGGAACCTGCGGGTATTTCTGTCATTGAGGAGACACCAT 2220
Db	1270 GAGATGGAATTTGCTGACCCCTGGAACCTGCGGGTATTTCTGTCATTGAGGAGACACCAT 1329
Qy	2221 AGTGAACCTGGGCTTATGCTGCTCCAGGCTGTGAGTGGTGGGAGACCTGTGTC 2280
Db	1330 AGTGAACCTGGGCTTATGCTGCTCCAGGCTGTGAGTGGTGGGAGACCTGTGTC 1389
Qy	2281 AGGTGTGATAGCCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTC 2339
Db	1390 AGGTGTGATAGCCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTC 1448

RESULT 11	2129 bp	mRNA	linear	ROD 16-APR-2003
BC033277	Mus musculus	RIKEN cDNA 2600017P15 gene, mRNA	(cdna clone MGC:28694	
LOCUS	IMAGE:4241041), complete cds.			
DEFINITION	BC033277.1 GI:23958794			
ACCESSION	MGC.			
VERSION	Mus musculus (house mouse)			
KEYWORDS	Mus musculus			
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
ORGANISM	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 2129)			
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Platchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.			
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			

TNGLIANIFTVAKTEVYVDSGSKDKMTAFIVERDFGINTGKPEKDLGIRGNTCE
VFHENTRVPVLEGEVGGFKVAMNINLSGRFSMGSAGVAGMLKLLIELTAETARK
QPNRLNLSFALIOELGSGVYKDYPIERMILDRARILLIFEGNEILRFLTALQAGHR
AAWOCVSEALIOELGSGVYKDYPIERMILDRARILLIFEGNEILRFLTALQAGHR
ILTSRKLKSGNVTYMETIGRKLRSLSLRTDGLGVVHPSLGSANKLEH
VHYFRTVETLLLRFGKNIVEQVLKRVANILINLYGMTAVLSRASRSIRIGLRNH
HEVLIANMFCVEAYFQNLFSLSQDKNAPENLDQIKKVSRIQLEKRAYICAPLDR
S"

BASE COUNT 554 a 504 c 593 g 497 t
ORIGIN

Query Match 56.6%; Score 1388.4; DB 10; Length 2148;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY	98	CGGCTGCGGCTCGTCCGCTGCGGCTGCTGGTGTCTACCGGCAACCGCGGCTACTGC	157
DB	94	CGGCTGCGGCGGAGCTCCCGGCTGCTGAGGGAATTCACGGGAGCGCGCGCGCTGC	153
QY	158	GCACGAGCCGCTGTAGGAGCTTTCGCCAAAGAGCTTTCTAGGCAAAATCAAGAAGA	217
DB	154	ACACGAGCTTCAGAGCTGCTGCTTCGCCAAGAGCTTCTTCCTGGGCAACATCAAGCAGA	213
QY	218	AAGAAGTTTCCCATTTCCAGAGCTTAGCCAAAGTGAACCTTAATGAATCAATCAGTTCT	277
DB	214	AAGAGTTTCCCATTTCCAGAGCTCAGTCAACATGAACCTTAGTGAATCAATCAGTTCT	273
QY	278	TGGAGCCGCTGGAAAAATTTCTCACTGAAGAGGTGGACTCCCGCAAAATTTGACCAAGAAG	337
DB	274	TGGAGCCGCTGGAAAAATTTCTCACTGAAGAGGTGGACTCCCGCAAAATTTGACCAAGAAG	333
QY	338	GAAAAATCCAGATGAACCTTTGGAGAAATTTGAAGAGCTTAGGGCTTTTGGGCTGCAAG	397
DB	334	GAANAATCCAGTTTGACACCTTAGAGAAGTTGAAGAGCTTGGGACTTTTGGCATACAGG	393
QY	398	TCCAGAGAATATGTGGCTGGGCTTCTCCTCAACACCATGTAATCAAGACTAGGGGAGA	457
DB	394	TCCAGAGAATATGTGGCTGGGCTTCTCCTCAACACCATGTAATCAAGACTAGGGGAGA	453
QY	458	TCATCAGCATGGATGGGTTCATCAGTGTGACCTGGCAGCGCACCAAGGCTATTTGGGCTCA	517
DB	454	TCATCAGCTGGATGGGTTCATCAGTGTGACCTGGCAGCGCACCAAGGCTATTTGGGCTCA	513
QY	518	AGGGATCATCTTTGGCTGGGCTGAGGAGCAGAAAGCCAAATACCTTGCCTAACTGGGCT	577
DB	514	AGGGATCATCTTTGGCTGGGCTGAGGAGCAGAAAGCCAAATACCTTGCCTAACTGGGCT	573
QY	578	CCGGGAGACATTTGAGGCTTCTGCTCACGGAGCCAGCCAGTGGGAGCGATGACGCT	637
DB	574	CTGGGAGACATTTGAGGCTTCTGCTCACGGAGCCAGCCAGTGGGAGTGTGCTGAT	633
QY	638	CAATCCGGAGCAGCCACACTAAGTGAAGACAGAAGCACTACATCTCAATGGCTCCA	697
DB	634	CAATCCAGACAGAGCTAGTGAAGTGAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	693
QY	698	AGGCTGGATTACTAATGAGGACTGGCCATATTTTACTGTGTTTCCAAAGACTGAGG	757
DB	694	AGGTTGGATCACCAGTGGGAGCTGGCCATATTTTACTGTGTTTCCAAAGACTGAGG	753
QY	758	TGCTTGATTCGATGGATCAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACT	817
DB	754	TGCTTGATTCGATGGATCAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACT	813
QY	818	TTGGTGGAGTCACTAATGGAAACCCGAGATTAATTAGGCATTCGGGCTTCCCAACTT	877
DB	814	TCGGCGGATCACTAATGGAAACCCGAGATTAATTAGGCATTCGGGCTTCCCAACTT	873
QY	878	GTGAAGTCCATTTTGAAGAACCCAGATACCTGTGGAAACATCTTTGGAGAGTCCGAG	937
DB	874	GTGAGTCCATTTTGAAGAACCCAGATACCTGTGGAAACATCTTTGGAGAGTCCGAG	933
QY	938	ATGGGTTTAAAGTGGCCATGAACATCTTCAACAGCGCGGCTTACAGCATGGGAGCGTCG	997
DB			

DB	934	GGGCTTTTAGGTAGCCATGAACATCTTGAACAGTGGCGGATTTACCATGGGCGTCTG	993
QY	998	TGCTGGCTGGCTCAAGAGATGATTAATGACCTGCTGAGTACGGCTGCACAGAAGAAC	1057
DB	994	TGCGCGGGATGCTGAAGAAGCTGATGCAATGACTGCTGAGTATGCTGTACGAGGAAC	1053
QY	1058	AGTTTAAACAGAGGCTCAGTGAATTTGGATTTGATTTCAGGAGAAATTTGGCAGTGGCTC	1117
DB	1054	AGTTTAAACAGAGATCTCAGTGAATTTGGGTTGATTTCAGGAAAGTTTGGCTCATGGCTC	1113
QY	1118	AGAAGCTTACCTCATGGAGAGATGACCTACCTCAGACGAGGAGTGTGGACCAACCTG	1177
DB	1114	AGAAGCATACCTCATGGAGAGATGAGGCTTACCTCAGTTCAGGAGTGTAGACCAACCG	1173
QY	1178	GCTTTCCGACTGCTCCATTCAGAGGAGGATGCTGAGGAGTGTGTTAGCTTCCGAGCGGCT	1237
DB	1174	GATTTCCGACTGCTTATTCAGAGGAGGATGAGGAGTGTGTTAGCTTCCGAGCGGCT	1233
QY	1238	GCGAGTGTGAGTGAAGGCTGAGAGTCTCGGGGGCTTGGGCTACACAGGAGCTATC	1297
DB	1234	GCGAGTGTGAGGAGGCTTCTGAGAGTCTTGGGGGCTCAGGCTTACATGAAGGACTACC	1293
QY	1298	CGTACGAGCGCATCTGCTGACACCCGCTCCTCCTCATCTTTCAGGAGGAGCAATGAGA	1357
DB	1294	CCTATGAGCGCATCTGCTGATGCCGCTTCTCCTTAATCTTTGAGGAGGAGCAATGAGA	1353
QY	1358	TTCTCCGATGTACATCCCTGAGGGGTCTGAGCATGCGGGGCTGAGCTGACTACCA	1417
DB	1354	TTCTCCGATGTACATCCCTGAGGGGTCTGAGCATGCGGGGCTGAGCTGACTACCA	1413
QY	1418	GGATCCATGAGCTTAAACAGGCGCAAGTGAAGCAGATCATGATGATGATGATGATGATG	1477
DB	1414	GGATCCATGAGCTTAAACAGGCGCAAGTGAAGCAGATCATGATGATGATGATGATGATG	1473
QY	1478	TTGGGACTTCTCGGGCGGAACTGTGGACCTTGGGGCTGACAGGCAACCATGAGTGTGCTG	1537
DB	1474	TTGGGACTTCTCGGGCGGAACTGTGGACCTTGGGGCTGACAGGCAACCATGAGTGTGCTG	1533
QY	1538	ACCCAGCTTTCGGGAGCAGTCCCAAGAGTTTGAAGAGAACACTTACTGCTCGGCGGGA	1597
DB	1534	ACCCAGCTTTCGGGAGCAGTCCCAAGAGTTTGAAGAGAACACTTACTGCTCGGCGGGA	1593
QY	1598	CCGTGGAGACACTGCTGCTGCTTGGCAAGCAGCATGAGGAGGAGCACTGCTGCTGCTG	1657
DB	1594	CTGTGGAGACCTGCTGCTGCTTGGCAAGCAGCATGAGGAGGAGCACTGCTGCTGCTG	1653
QY	1658	AGCGGTGGGCAACATCTCATCAACCTGTATGGCATGACGGCGCTGCTGCTGCTGCTG	1717
DB	1654	AGCGGTGGGCAACATCTCATCAACCTGTATGGCATGACGGCGCTGCTGCTGCTGCTG	1713
QY	1718	GCGCTCCATCCGCTTGGGCTCCGCAACCCAGCAGGAGTCTTCTTGGGCAACACTT	1777
DB	1714	GCGCTCCATCCGCTTGGGCTCCGCAACCCAGCAGGAGTCTTCTTGGGCAACACTT	1773
QY	1778	TCTGCTGGAGGCTTACTTGCAGAACTCTTTCAGGCTCTCTCAGCTGGACAACTGCTG	1837
DB	1774	TCTGCTGGAGGCTTACTTGCAGAACTCTTTCAGGCTCTCTCAGCTGGACAACTGCTG	1833
QY	1838	CAGAAACCTTAGATGAGCAGATTAAGAAAGTGTCCAGCAGATCTTGTAGAGAGCGGCT	1897
DB	1834	CAGAAACCTTAGATGAGCAGATTAAGAAAGTGTCTCGCAGATCTTGTAGAGAGCGGCT	1893
QY	1898	ATATCTGTGCCACCTCTGGCAGGAGCATGCTGAGGAGGGGACA	1943
DB	1894	ATATCTGTGCCACCTCTGGCAGGAGCATGCTGAGGAGGGGACA	1939

RESULT 13
BC031137
LOCUS
DEFINITION
IMAGE:4971848, complete cds.
ACCESSION
BC031137

BC031137 3071 bp mRNA linear ROD 10-JUN-2003
Mus musculus RIKEN cDNA 2600017P15 gene, mRNA (cdna clone MGC:37218
IMAGE:4971848), complete cds.

1805 TCTTCAGCCTCTCTCAGCTGGCACAAGTATGCTCCAGAAACCTAGATGAGCAGATTAA 1864
1363 TCTTCAGCCTCTCTCAGCTGGCACAAGTATGCTCCAGAAACCTAGATGAGCAGATTAA 1422

FEATURES		Location/Qualifiers		/organism='Homo sapiens (human)'	
source		1..771		/organism='Homo sapiens'	
				/mol_type='genomic DNA'	
				/db_xref='taxon:9606'	
BASE COUNT		204 a 187 c 219 g 158 t		3 others	
ORIGIN					
Query Match		30.6%; Score 750.6; DB 6; Length 771;			
Best Local Similarity		99.08; Pred. No. 2.6e-175;			
Matches 764; Conservative 0; Mismatches 7; Indels 1; Gaps 1;					
QY	21	GCTAAGAGGGGAGACTGAGCGCTGGGGAACATCGGGCAGCATGAGCGCTGCGG	80		
Db	1	GCTAAGAGGGGAGACTGAGCGCTGGGGAACATCGGGCAGCATGAGCGCTGCGG	60		
QY	81	GCTCTTCTCGCGCACCGCTGGGCTCGCTGCGCGGGGTCTGCTGTCTTACCGC	140		
Db	61	GCTCTTCTCGCGCACCGCTGGGCTCGCTGCGCGGGGTCTGCTGTCTTACCGC	120		
QY	141	GAACCGCGGCTACTGCGCACCGCGCTGTACGAGCTTCCGCAAGAGCTTTTCCT	200		
Db	121	GAACCGCGGCTACTGCGCACCGCGCTGTACGAGCTTCCGCAAGAGCTTTTCCT	180		
QY	201	AGGCAAAATCAAGAGAGAGAGATTTTCCCAATTTCCAGAACTTAGCCAAAGAGCTTAA	260		
Db	181	AGGCAAAATCAAGAGAGAGAGATTTTCCCAATTTCCAGAACTTAGCCAAAGAGCTTAA	240		
QY	261	TCAATCAATCAGTTCTTTGGGACCGCTGGGAAATTTCTTCACTGAAGAGTGGACTCCG	320		
Db	241	TCAATCAATCAGTTCTTTGGGACCGCTGGGAAATTTCTTCACTGAAGAGTGGACTCCG	300		
QY	321	AAAAATTGACAGAGAGAGAGATTTCCAGATGCAACTTTGGAGAAATTTGAAGAGCTAGG	380		
Db	301	AAAAATTGACAGAGAGAGAGATTTCCAGATGCAACTTTGGAGAAATTTGAAGAGCTAGG	360		
QY	381	GCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCGCTGGGCTTCTCCACACCATGTA	440		
Db	361	GCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCGCTGGGCTTCTCCACACCATGTA	420		
QY	441	CTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCACTGACCTGGCAGCGCA	500		
Db	421	CTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCACTGACCTGGCAGCGCA	480		
QY	501	CCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCCTGAGGAGCAAGCAATA	560		
Db	481	CCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCCTGAGGAGCAAGCAATA	540		
QY	561	CTTGCTAACTGGGCTCGGGGAGACATTCAGCCCTTCTGCTCAGGAGCCAGCCAG	620		
Db	541	CTTGCTAACTGGGCTCGGGGAGACATTCAGCCCTTCTGCTCAGGAGCCAGCCAG	600		
QY	621	TGGGAGGATGAGCGCTCAATCCGAGCAGAGCCACACTAAGTGAAGCAAGAACACTA	680		
Db	601	TGGGAGGATGAGCGCTCAATCCGAGCAGAGCCACACTAAGTGAAGCAAGAACACTA	660		
QY	681	CATCTCTCAATGGCTCCCAAGGCTCTGATTAATGAGGAGCTGGCCAAATTTTACTCT	740		
Db	661	CATCTCTCAATGGCTCCCAAGGCTCTGATTAATGAGGAGCTGGCCAAATTTTACTCT	720		
QY	741	GTTTGGCAAGACTGAGGCTGTTGATTTCTGATGGATCAGTGAAGACAAATC	792		
Db	721	GTTTGGCAAGACTGAGGCTGTTGATTTCTGATGGATCAGTGAAGACAAATC	771		

Search completed: September 6, 2003, 17:23:04
Job time : 9029.27 secs

QY	1865	AAAGTCTCCAGCAGATCCTTGAGAGCAGCGCTATATCTGTGCCACCCCTCTGCACAGGA	1924		
Db	1423	AAAGTCTCCAGCAGATCCTTGAGAGCAGCGCTATATCTGTGCCACCCCTCTGCACAGGA	1482		
QY	1925	CATCTGAGGAGGAGAGTGTCCCTGCTACC--GCCCGCCCTACCCATGGCC--GT	1981		
Db	1483	CATCTGAGGAGGAGAGTGTCCCTGCTACC--GCCCGCCCTACCCATGGCC--GT	1542		
QY	1982	TGCTGGATGACTGTACTCTTTTTCAGAAGGTGTGGATATACAGAGTTAAGCCCTT	2041		
Db	1543	TGCTGGATGACTGTACTCTTTTTCAGAAGGTGTGGATATACAGAGTTAAGCCCTT	1602		
QY	2042	TGTTCCCGCTGTCACCTGAAGGTGTGCTGCGCTGGGAGAGCCCTTCCAGGTTT	2101		
Db	1603	TGTTCCCGCTGTCACCTGAAGGTGTGCTGCGCTGGGAGAGCCCTTCCAGGTTT	1662		
QY	2102	GACCTCGAGGAGTGTCTTCAACAGGACCATCACAGCTTCTGAAGTGGCGGAGAGAG	2161		
Db	1663	GACCTCGAGGAGTGTCTTCAACAGGACCATCACAGCTTCTGAAGTGGCGGAGAGAG	1722		
QY	2162	AGAAATGGAATGTGACCCCTGGAAGTGGGCTGATCTGATGAGGAGACACCATTA	2221		
Db	1723	AGAAATGGAATGTGACCCCTGGAAGTGGGCTGATCTGATGAGGAGACACCATTA	1782		
QY	2222	GTGAAACTGGGCTTATGCTGCTGCTCCAGGCTGTGAGTGGGTGGGACCTGTGTCA	2281		
Db	1783	GTGAAACTGGGCTTATGCTGCTGCTCCAGGCTGTGAGTGGGTGGGACCTGTGTCA	1842		
QY	2282	GGTGTGGATAGCCATTTGCTCAACACACATCTCTAAGAACAGCTTGAAGCTCG	2341		
Db	1843	GGTGTGGATAGCCATTTGCTCAACACACATCTCTAAGAACAGCTTGAAGCTCG	1902		
QY	2342	TCTGGTCAATCAATTAAGTGAAGCAGAGGACATTAAGCAATGTACCAAGAACCATTT	2401		
Db	1903	TCTGGTCAATCAATTAAGTGAAGCAGAGGACATTAAGCAATGTACCAAGAACCATTT	1962		
QY	2402	AACAAAGATATAAATGTCCACAA	2425		
Db	1963	AACAAAGATATAAATGTCCACAA	1986		
RESULT 15					
BD150623					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
1 (bases 1 to 771)					
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,					
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.					
Primer for synthesizing full-length cDNA and use thereof					
Patent: JP 2002191363-A 5466 09-JUL-2002;					
HELEX RESEARCH INSTITUTE					
OS Homo sapiens (human)					
PN JP 2002191363-A/5466					
PD 09-JUL-2002					
PF 28-JUL-2000 JP 2000280990					
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU					
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,					
PI KEIICHI NAGAI, TETSUJI OTSUKI					
PC					
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10,					
PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC					
Primer for synthesizing full-length cDNA and use thereof FH key					
Location/Qualifiers					
1..771					
FT source					

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	2452	100.0	2452	24	AAD34006		Human acyl dehydro
2	2452	100.0	2522	23	ABV21123		Human prostate exp
3	2452	100.0	2522	23	ABV23498		Human prostate exp
4	2452	100.0	2522	23	ABV26965		Human prostate exp
5	2452	100.0	2522	23	ABV29355		Human prostate exp
6	2450.6	99.9	2492	22	AAS40878		cDNA encoding nove
7	2438.4	99.4	2440	22	AH24246		Human oxidoreducta
8	2382.2	97.2	2399	22	AAH16781		Human cDNA sequenc

FT FT /note= "This region is specifically claimed in
FT claim 1 as SEQ ID NO:3"
FT 1933..2452
FT /*tag= C
XX
PN WO200218582-A2.
XX 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US27186.
XX 31-AUG-2000; 2000US-229831P.
PR (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers R, Hunter JJ;
XX WPI; 2002-329775/36.
DR P-PSDB; AAE21680.
XX
XX New human dehydrogenase polypeptide for diagnosing and treating
PT dehydrogenase-7 associated diseases and disorders e.g. Alzheimer's
PT disease, and to identify modulators of therapeutic use
XX
XX Claim 1; Fig 1; 124pp; English.
XX
XX The invention relates to human dehydrogenase (DHDR)-7 polypeptides
CC referred to as 62112 and nucleic acid molecules encoding such
CC polypeptides. DHDR-7 is an acyl-CoA dehydrogenase, a mitochondrial
CC flavoprotein which catalyses the first step of fatty acid beta-
CC oxidation. Sequences of the invention and their antibodies are
CC useful for treating a disorder, disease or condition which is caused
CC by misregulation (e.g. downregulation or upregulation) of DHDR activity.
CC Examples of disorders include central nervous system (CNS) disorders
CC e.g. Alzheimer's disease (AS), Pick's disease, Parkinson's, Huntington's
CC disease, Gilles de la Tourette's syndrome, multiple sclerosis, epilepsy,
CC amyotrophic lateral sclerosis and Creutzfeldt-Jakob disease; autonomic
CC function disorders e.g. hypertension, depression, schizophrenia, panic
CC migraine, anxiety, obsessive-compulsive disorder and obesity; cardiac
CC related disorder e.g. arteriosclerosis, ischaemia reperfusion injury,
CC restenosis, arterial inflammation, angina, hypertension, cardiomyopathy
CC and arrhythmia; disorders of musculus skeletal system paralysis, muscle
CC weakness e.g. ataxia; cell proliferation, growth, differentiation or
CC migration disorders e.g. cancer, diabetes mellitus, hypothyroidism or
CC hyperthyroidism; reproductive or fertility disorders; autoimmune or
CC or immune deficiency disorders; hepatic disease or dysfunction and
CC metabolic disorders. They are used for screening assays, predictive
CC medicine e.g. diagnostic assays, prognostic assays, monitoring clinical
CC trials, and pharmacogenetics. Polypeptides of the invention are used
CC to identify modulators that modulate their activity. Polynucleotides
CC of the invention are used in gene therapy. The present sequence is
CC cDNA encoding human DHDR-7.
XX
XX Sequence 2452 BP; 616 A; 616 C; 685 G; 535 T; 0 other;
SQ

Query Match 100.0%; Score 2452; DB 24; Length 2452;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQTGTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCG 60
DB 1 CQTGTGTGTGTCCTCGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCG 60
QY 61 GGCAGCATGAGCGGTGCGGGCTCTTCCTGCGCACCGAGGCTGCGGCTGCGGCTGCGG 120
DB 61 GGCAGCATGAGCGGTGCGGGCTCTTCCTGCGCACCGAGGCTGCGGCTGCGGCTGCGG 120
QY 121 GGTCTGTGTGTCCTACCGCAACCGCGGCTACTGCGCACCGAGCCCGCTGTACGAGCT 180
DB 121 GGTCTGTGTGTCCTACCGCAACCGCGGCTACTGCGCACCGAGCCCGCTGTACGAGCT 180
QY 181 TTCGCCAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAGTTTCCCATTTCCAGAA 240
DB 181 TTCGCCAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAGTTTCCCATTTCCAGAA 240

181 TTCGCCAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAGTTTCCCATTTCCAGAA 240
QY 241 GTTAGCCCAAGCTTAATGAAATCAATCACTTCTTGGACCCGCTGGAAAATTTCTTC 300
DB 241 GTTAGCCCAAGCTTAATGAAATCAATCACTTCTTGGACCCGCTGGAAAATTTCTTC 300
QY 301 ACTCAAGAGGTGGATCCCGAAAAATGACCAAGAGGGAATCCAGATGAACCTTG 360
DB 301 ACTCAAGAGGTGGATCCCGAAAAATGACCAAGAGGGAATCCAGATGAACCTTG 360
QY 361 GAGAAATGAGAGCCCTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGCTGGCCTG 420
DB 361 GAGAAATGAGAGCCCTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGCTGGCCTG 420
QY 421 GGCCTTCCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGATGGATGGTCCATC 480
DB 421 GGCCTTCCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGATGGATGGTCCATC 480
QY 481 ACTGTGACCTGGCAGCGCCAGGCTATTGGCCCTCAAGGGGATCATCTTTGGCTGGCCT 540
DB 481 ACTGTGACCTGGCAGCGCCAGGCTATTGGCCCTCAAGGGGATCATCTTTGGCTGGCCT 540
QY 541 GAGGAGCAAGAAAGCAAAATCTTGCCTAAACTGCGCTCCGGGGAGACATTTGAGCCTTC 600
DB 541 GAGGAGCAAGAAAGCAAAATCTTGCCTAAACTGCGCTCCGGGGAGACATTTGAGCCTTC 600
QY 601 TGCCTTACGGAGCCAGCCAGTGGGAGCGATCGACCTCAATCCGGAGCAGAGCCACACTA 660
DB 601 TGCCTTACGGAGCCAGCCAGTGGGAGCGATCGACCTCAATCCGGAGCAGAGCCACACTA 660
QY 661 AGTGAAGACAAGAACACTACATCTCAATGGCTCCAAAGTCTCGATTAATGGAGGA 720
DB 661 AGTGAAGACAAGAACACTACATCTCAATGGCTCCAAAGTCTCGATTAATGGAGGA 720
QY 721 CTGGCCAAATATTTTACTGTGTTCGCAAGACTGAGGCTGTTGATCTGATGGATCAGTG 780
DB 721 CTGGCCAAATATTTTACTGTGTTCGCAAGACTGAGGCTGTTGATCTGATGGATCAGTG 780
QY 781 AAAGCAAAATCAAGAGCACTATAGTAAAGAGACTTTGGTGGAGTCACTAATGGGAAA 840
DB 781 AAAGCAAAATCAAGAGCACTATAGTAAAGAGACTTTGGTGGAGTCACTAATGGGAAA 840
QY 841 CCCGAAGATAAATTAGGCAATCGGGCTCCAAACACTTGTGAAGTCCATTTGAAACACC 900
DB 841 CCCGAAGATAAATTAGGCAATCGGGCTCCAAACACTTGTGAAGTCCATTTGAAACACC 900
QY 901 AGATACCTGTGGAACAACTCTTGGAGAGTTCGAGATGGGTTTAAAGTGCCCATGAAC 960
DB 901 AGATACCTGTGGAACAACTCTTGGAGAGTTCGAGATGGGTTTAAAGTGCCCATGAAC 960
QY 961 ATCCTCAACAGCGCGGTTTCAGCATGGGCGCTGCTGGGCTGCTCAAGAGATTG 1020
DB 961 ATCCTCAACAGCGCGGTTTCAGCATGGGCGCTGCTGGGCTGCTCAAGAGATTG 1020
QY 1021 ATTGAATGACTGCTGAGTACGCTGCACAAGAAACACTTTTAAAGAGGCTCAGTGAA 1080
DB 1021 ATTGAATGACTGCTGAGTACGCTGCACAAGAAACACTTTTAAAGAGGCTCAGTGAA 1080
QY 1081 TTTGGATTGATTTCAGAGAAATTTGCACCTGATGGCTCAGAAGGCTTACGTCATGAGAGT 1140
DB 1081 TTTGGATTGATTTCAGAGAAATTTGCACCTGATGGCTCAGAAGGCTTACGTCATGAGAGT 1140
QY 1141 ATGACCTTACCTACAGCAGGGATGCTGGACCAACCTTGGCTTCCGACCTGCTCCATCGAG 1200
DB 1141 ATGACCTTACCTACAGCAGGGATGCTGGACCAACCTTGGCTTCCGACCTGCTCCATCGAG 1200
QY 1201 GCAGCCATGCTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGGCGCTG 1260
DB 1201 GCAGCCATGCTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGGCGCTG 1260
QY 1261 CAGATCTCGGGGCTTGGGCTACAAAGGAGTATCCGTACAGCGGCTACTGCGCTGAC 1320
DB 1261 CAGATCTCGGGGCTTGGGCTACAAAGGAGTATCCGTACAGCGGCTACTGCGCTGAC 1320

Db	111	GGCAGCATGAGCGGCTGCGGGCTCTCCCTGCGCACCACGGCTGCGGCTGCTGCGCTG	170
Qy	121	GGTCTGGTGTCTCTACCGCGAACCGCGGGCTACTGCGGCACACGGCCCGCTGTACGAGCT	180
Db	171	GGTCTGGTGTCTCTACCGCGAACCGCGGGCTACTGCGGCACACGGCCCGCTGTACGAGCT	230
Qy	181	TTCCGCCAAGAGAGCTTTTCCCTAGCAGAAATCAGAGAGAAAGATTTCCTCCATTTCCAGAA	240
Db	231	TTCCGCCAAGAGAGCTTTTCCCTAGCAGAAATCAGAGAGAAAGATTTCCTCCATTTCCAGAA	290
Qy	241	GTTAGCACAAGTAACTTAATGAATCAATCACTTCTGGGACCCCGTGAAATAATTTCTTC	300
Db	291	GTTAGCACAAGTAACTTAATGAATCAATCACTTCTGGGACCCCGTGAAATAATTTCTTC	350
Qy	301	ACTGAGAGGTGGACTCCCGAAATAATTGACGAGGAAGGAAATCCACATGAACTTTG	360
Db	351	ACTGAGAGGTGGACTCCCGAAATAATTGACGAGGAAGGAAATCCACATGAACTTTG	410
Qy	361	GAGAAATGAGAGCCTAGGCTTTTGGGCTCAAGTCCAGAGAAATATGCTGGCCTG	420
Db	411	GAGAAATGAGAGCCTAGGCTTTTGGGCTCAAGTCCAGAGAAATATGCTGGCCTG	470
Qy	421	GGCTTCTCCAAACCACTACTCAAGACTAGGGGAGATCATAGCATGGATGGGTCATC	480
Db	471	GGCTTCTCCAAACCACTACTCAAGACTAGGGGAGATCATAGCATGGATGGGTCATC	530
Qy	481	ACTGTGACCTGGCAGCGCACCAGGCTATTGGGCTCAAGGGGATCATCTTTGGCTGGCACT	540
Db	531	ACTGTGACCTGGCAGCGCACCAGGCTATTGGGCTCAAGGGGATCATCTTTGGCTGGCACT	590
Qy	541	GAGGAGCAAGAACCAATACTTGCCTAACTGGCGTCCGGGGAGGACATTTGACGCGTTC	600
Db	591	GAGGAGCAAGAACCAATACTTGCCTAACTGGCGTCCGGGGAGGACATTTGACGCGTTC	650
Qy	601	TGCCTCAGGAGCAGCCAGTGGGAGGATCGACGCTCAATCCGAGCAGGACACACTA	660
Db	651	TGCCTCAGGAGCAGCCAGTGGGAGGATCGACGCTCAATCCGAGCAGGACACACTA	710
Qy	661	AGTGAAGACAAGACACTACATCCTCAATGGCTCCAGGTCCTGGATTAATTTGAGGA	720
Db	711	AGTGAAGACAAGACACTACATCCTCAATGGCTCCAGGTCCTGGATTAATTTGAGGA	770
Qy	721	CTGGCCAAATTTTACTGTGTTGCCAAGACTGAGGTCTGTGATCTGTGATGATCAGTG	780
Db	771	CTGGCCAAATTTTACTGTGTTGCCAAGACTGAGGTCTGTGATCTGTGATGATCAGTG	830
Qy	781	AAAGACAATAACACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTTAATGGGAAA	840
Db	831	AAAGACAATAACACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTTAATGGGAAA	890
Qy	841	CCCGAAGATAAATTAGGCATTCGGGGCTCCAACTGTGTGAAGTCCATTTTGAAGACAC	900
Db	891	CCCGAAGATAAATTAGGCATTCGGGGCTCCAACTGTGTGAAGTCCATTTTGAAGACAC	950
Qy	901	AGATACCTGTGGAAACATCCTTGSAGAGTCCGAGATGGGTTTAAAGTGGCCATGAAC	960
Db	951	AGATACCTGTGGAAACATCCTTGSAGAGTCCGAGATGGGTTTAAAGTGGCCATGAAC	1010
Qy	961	ATCCTCAACAGCGCGGTTACAGCATGGGACGCTGCGTGGGCTGCTCAAGAGATTG	1020
Db	1011	ATCCTCAACAGCGCGGTTACAGCATGGGACGCTGCGTGGGCTGCTCAAGAGATTG	1070
Qy	1021	ATTGAATGACTGCTGAGTACGCTCCACAGGAACAGTTTAAAGAGGCTCAGTGAA	1080
Db	1071	ATTGAATGACTGCTGAGTACGCTCCACAGGAACAGTTTAAAGAGGCTCAGTGAA	1130
Qy	1081	TTTGGATTGATTGAGGAGAAATTTGACATGATGGCTCAGAGGCTTAACATGAGAGT	1140
Db	1131	TTTGGATTGATTGAGGAGAAATTTGACATGATGGCTCAGAGGCTTAACATGAGAGT	1190
Qy	1141	ATGACCTACCTACAGCAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAG	1200
Db	1191	ATGACCTACCTACAGCAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAG	1250
Qy	1201	GCAGCCATGGTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGGAGCGCTG	1260
Db	1251	GCAGCCATGGTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGGAGCGCTG	1310
Qy	1261	CAGATCTCTGGGGCTTGGGCTACACAAGGAGACTATCCGTACGAGCGCATCTGCGTGAC	1320
Db	1311	CAGATCTCTGGGGCTTGGGCTACACAAGGAGACTATCCGTACGAGCGCATCTGCGTGAC	1370
Qy	1321	ACCGCATCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTG	1380
Db	1371	ACCGCATCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTG	1430
Qy	1381	ACGGCTCTGACAGCATGCGCGCGCATCTCTGACTACCAGGATCCATGAGCTTAAACAGGCC	1440
Db	1431	ACGGCTCTGACAGCATGCGCGCGCATCTCTGACTACCAGGATCCATGAGCTTAAACAGGCC	1490
Qy	1441	AAAGTGACACAGTATGATACCTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAAC	1500
Db	1491	AAAGTGACACAGTATGATACCTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAAC	1550
Qy	1501	GTGGACCTGGGCTGACAGGCAACCATGGAGTTGTGACCCCGAGTCTTTCGGGACAGTGCC	1560
Db	1551	GTGGACCTGGGCTGACAGGCAACCATGGAGTTGTGACCCCGAGTCTTTCGGGACAGTGCC	1610
Qy	1561	AAACAAGTTGAGGAGAACACACTACTGCTTCGSCCGGACCGTGGAGACACTGCTGCTCGC	1620
Db	1611	AAACAAGTTGAGGAGAACACACTACTGCTTCGSCCGGACCGTGGAGACACTGCTGCTCGC	1670
Qy	1621	TTTGGCAAGACCATCATGAGGAGGAGCTGTGACTTGAAGCGGCTGGCCAAACATCTCATC	1680
Db	1671	TTTGGCAAGACCATCATGAGGAGGAGCTGTGACTTGAAGCGGCTGGCCAAACATCTCATC	1730
Qy	1681	AACCTGTATGAGATGAGCGCGTGTGTCGCGGGCCAGCGCTCCATCCGATTTGGGCTC	1740
Db	1731	AACCTGTATGAGATGAGCGCGTGTGTCGCGGGCCAGCGCTCCATCCGATTTGGGCTC	1790
Qy	1741	CGCAACACGACACACAGAGTTCTCTTGGGCAACACCTTCTGCGTGGAAAGCTTACTTGCAG	1800
Db	1791	CGCAACACGACACACAGAGTTCTCTTGGGCAACACCTTCTGCGTGGAAAGCTTACTTGCAG	1850
Qy	1801	AATCTCTTACGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATT	1860
Db	1851	AATCTCTTACGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATT	1910
Qy	1861	AAGAAAGTGTCCAGCAGATCCCTTGAGAGCGAGCCTATATCTGTGCCCAACCTCTGGAC	1920
Db	1911	AAGAAAGTGTCCAGCAGATCCCTTGAGAGCGAGCCTATATCTGTGCCCAACCTCTGGAC	1970
Qy	1921	AGGACATGTGAGCGAGGAGACAGTGTCCCTCTACCGCCCGCCCTACCCATGGCCCG	1980
Db	1971	AGGACATGTGAGCGAGGAGACAGTGTCCCTCTACCGCCCGCCCTACCCATGGCCCG	2030
Qy	1981	TTGCTGATCAGCTGTACTCTTTTTCAGAAAGTGTGGGATTTATCACAGTTTAAAGCTTT	2040
Db	2031	TTGCTGATCAGCTGTACTCTTTTTCAGAAAGTGTGGGATTTATCACAGTTTAAAGCTTT	2090
Qy	2041	TTGTTTCCCGCTGTGACCTGAAAGGTTGTCCCTGGGCTGGGAGAGGCTCTTCCAGGTTT	2100
Db	2091	TTGTTTCCCGCTGTGACCTGAAAGGTTGTCCCTGGGCTGGGAGAGGCTCTTCCAGGTTT	2150
Qy	2101	TGACCTGGAGCAGTGTCTCTACAGGACCATCACAGTCTTGAACCTGAGCCGGAGAGA	2160
Db	2151	TGACCTGGAGCAGTGTCTCTACAGGACCATCACAGTCTTGAACCTGAGCCGGAGAGA	2210
Qy	2161	GAGAAATGAAATGCTGACCCCTGGAACCTGGCGGATTTCTGGTCAATTTGAGGAGACCAT	2220
Db	2211	GAGAAATGAAATGCTGACCCCTGGAACCTGGCGGATTTCTGGTCAATTTGAGGAGACCAT	2270
Qy	2221	AGTGAAGACTGGGCTTATGCTGCTCCAGGGTGTGAGGTGGGAGGAGCTGTGTC	2280
Db	2271	AGTGAAGACTGGGCTTATGCTGCTCCAGGGTGTGAGGTGGGAGGAGCTGTGTC	2330

1081 TTTGGATTGATTCAGGAGAAATTTGCATGTATGGCTCAGAGGCTTACGTATGAGAGT 1140
1131 TTTGGATTGATTCAGGAGAAATTTGCATGTATGGCTCAGAGGCTTACGTATGAGAGT 1190
1141 ATGACCTACTCACAGCAGGATGCTGGACCAACCTGGCTTTCCGAGTGTCTCCATCGAG 1200
1191 ATGACCTACTCACAGCAGGATGCTGGACCAACCTGGCTTTCCGAGTGTCTCCATCGAG 1250
1201 GCAGCCATGTTGAAGTGTTCAGCTCCGAGGCGCTTGGCAGTGTGTGAGTGAGCGGTG 1260
1251 GCAGCCATGTTGAAGTGTTCAGCTCCGAGGCGCTTGGCAGTGTGTGAGTGAGCGGTG 1310
1261 CAGATCCTCGGGGCTTGGGCTACACAGAGGACTATCCGTAGACGGCATATCGTGTGAC 1320
1311 CAGATCCTCGGGGCTTGGGCTACACAGAGGACTATCCGTAGACGGCATATCGTGTGAC 1370
1321 ACCCGATCCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCCGCTG 1380
1371 ACCCGATCCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCCGCTG 1430
1381 ACCGGTCTGAGCATCGCGCGCATCTTCGAGGAGGCTTCCGAGTCTCCGCGCAACT 1500
1431 ACCGGTCTGAGCATCGCGCGCATCTTCGAGGAGGCTTCCGAGTCTCCGCGCAACT 1550
1441 AAGTGTAGCAGTGTGATGATACCTTGGCGGAGGCTTCCGAGTCTCCGCGCAACT 1560
1491 AAGTGTAGCAGTGTGATGATACCTTGGCGGAGGCTTCCGAGTCTCCGCGCAACT 1550
1501 GTGACCTGGGGCTGACAGGAAACCAATGAGATTGTGACCCAGTCTTCGGACAGTGC 1560
1551 GTGACCTGGGGCTGACAGGAAACCAATGAGATTGTGACCCAGTCTTCGGACAGTGC 1610
1561 AACAGTTTGGAGGAGACACTTCTGCTCGCGGAGGCTTCCGAGTCTTCGGACAGTGC 1620
1611 AACAGTTTGGAGGAGACACTTCTGCTCGCGGAGGCTTCCGAGTCTTCGGACAGTGC 1670
1621 TTTGGCAAGCAATCATGAGGAGGAGTGTGATGAGGAGGCTTCCGAGTCTTCGGACAGTGC 1680
1671 TTTGGCAAGCAATCATGAGGAGGAGTGTGATGAGGAGGCTTCCGAGTCTTCGGACAGTGC 1730
1681 AACCTGTATGGCATGAGCGCGCTGCTCGCGGAGGCTTCCGAGTCTTCGGACAGTGC 1740
1731 AACCTGTATGGCATGAGCGCGCTGCTCGCGGAGGCTTCCGAGTCTTCGGACAGTGC 1790
1741 CGCAACCAACCAACAGGTTCTCTGCGCAACCACTTCTGCGTGGAGGTTACTTCTGAG 1800
1791 CGCAACCAACCAACAGGTTCTCTGCGCAACCACTTCTGCGTGGAGGTTACTTCTGAG 1850
1801 AATCTCTTACGCTCTCAGCTGGACAAGTATGCTCCAGAAACCTAGATGAGCAGATT 1860
1851 AATCTCTTACGCTCTCAGCTGGACAAGTATGCTCCAGAAACCTAGATGAGCAGATT 1910
1861 AAGAAAGTGTCCAGCAGATCCTTTGAGAAGGAGGCTTATCTGTGCGCCACCTCTTGAC 1920
1911 AAGAAAGTGTCCAGCAGATCCTTTGAGAAGGAGGCTTATCTGTGCGCCACCTCTTGAC 1970
1921 AGGACATGCTGAGGAGGAGGAGTGTCCCTGCTACGCGCGCGCTTACCCATGCGCGG 1980
1971 AGGACATGCTGAGGAGGAGGAGTGTCCCTGCTACGCGCGCGCTTACCCATGCGCGG 2030
1981 TTGCTGGATGACTGTACTCTTTTTCAGAAGGTTGTTGGGATTTATCAGGTTAAGCTT 2040
2031 TTGCTGGATGACTGTACTCTTTTTCAGAAGGTTGTTGGGATTTATCAGGTTAAGCTT 2090
2041 TTGTTCCCGCTGTGACCTGAAGGTTGTGCGCTGGGCTGGGAGGCTTCTCAGGTTT 2100
2091 TTGTTCCCGCTGTGACCTGAAGGTTGTGCGCTGGGCTGGGAGGCTTCTCAGGTTT 2150
2101 TGACCTGAGGAGGAGTGTCTTACAGGACCATCACAGCTTCTGAACCTGAGCGCGGAGA 2160
2151 TGACCTGAGGAGGAGTGTCTTACAGGACCATCACAGCTTCTGAACCTGAGCGCGGAGA 2210

2161 GAGAAATGAATGCTGACCCCTGGAACCTGGGGGTATTTGTCATTGAGGAGACCAT 2220
2211 GAGAAATGAATGCTGACCCCTGGAACCTGGGGGTATTTGTCATTGAGGAGACCAT 2270
2221 AGTGAACACTGGGCTTATGCTGCTGCTCCAGGCTGTGAGGTGGGTGGGACTGTGTC 2280
2271 AGTGAACACTGGGCTTATGCTGCTGCTCCAGGCTGTGAGGTGGGTGGGACTGTGTC 2330
2281 AGTGTGGATAGCCATTTCTGCTCAACACACACATTTCTTAAGAAACAGCTTGAAGCTCT 2340
2331 AGTGTGGATAGCCATTTCTGCTCAACACACACATTTCTTAAGAAACAGCTTGAAGCTCT 2390
2341 GTCTGGGTCAATTAATAAGAGGAGGACATTAACACATGTACAGGAGACCAT 2400
2391 GTCTGGGTCAATTAATAAGAGGAGGACATTAACACATGTACAGGAGACCAT 2450
2401 TAACAAGAATATAAATGTACAACTCTGTCTACTGTATAAAAAA 2452
2451 TAACAAGAATATAAATGTACAACTCTGTCTACTGTATAAAAAA 2502

RESULT 4
ABV26965
ID ABV26965 standard; cDNA; 2522 BP.
XX
AC ABV26965;
XX
DT 16-SEP-2002 (first entry)
XX
Human prostate expression marker cDNA 26956.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
XX WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
XX Claim 1; Page 5464; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;

CC	(h) assessing the aggressiveness or indolence of prostate cancer in a patient;									
CC	(i) is also useful as a pharmacodynamic or pharmacogenomic marker.									
XX										
SQ	Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;									
	Query Match									
	Best Local Similarity 100.0%; Score 2452; DB 23; Length 2522;									
	Matches 2452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CCTGTGTGTGTCCTCCGCGCTAGAGAGGAGACTGAGGCTGAGGCTGGGGAACATCG	60							
DB	51	CCTGTGTGTGTCCTCCGCGCTAGAGAGGAGACTGAGGCTGAGGCTGGGGAACATCG	110							
QY	61	GGCAGCATGAGCGGTGCGGCTCTCTCTGCGCACCGGCTGAGGCTGGGCTGGGCTGGG	120							
DB	111	GGCAGCATGAGCGGTGCGGCTCTCTCTGCGCACCGGCTGAGGCTGGGCTGGGCTGGG	170							
QY	121	GCTCTGTGTGTCCTACCGCAACCGGCGGTACTGCGCACCGGCTGAGGCTGGGCTGGG	180							
DB	171	GCTCTGTGTGTCCTACCGCAACCGGCGGTACTGCGCACCGGCTGAGGCTGGGCTGGG	230							
QY	181	TTCCGCAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAAAGTTCCTGAGGCTGGG	240							
DB	231	TTCCGCAAGAGCTTTTCTAGGCAAAATCAAGAGAAAGAAAGTTCCTGAGGCTGGG	290							
QY	241	GTTAGCCAAAGATGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	300							
DB	291	GTTAGCCAAAGATGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	350							
QY	301	ACTCAAGAGTGGACTCCGCAAAATGACCAAGAGGAAATCCAGATGAATCAATCAAT	360							
DB	351	ACTCAAGAGTGGACTCCGCAAAATGACCAAGAGGAAATCCAGATGAATCAATCAAT	410							
QY	361	GAGAAATGAGAGCTTAGGCTTTTGGGCTGCAAGTCCAGAGAAATGAGGCTGGG	420							
DB	411	GAGAAATGAGAGCTTAGGCTTTTGGGCTGCAAGTCCAGAGAAATGAGGCTGGG	470							
QY	421	GGCTTCTCCCAACACCATGACTCAAGACTAGGAGGATCATCAGATGGGCTGGG	480							
DB	471	GGCTTCTCCCAACACCATGACTCAAGACTAGGAGGATCATCAGATGGGCTGGG	530							
QY	481	ACTGTGACCTGGCAGCCAGGCTATTTGGCTCAAGTCCAGAGAAATGAGGCTGGG	540							
DB	531	ACTGTGACCTGGCAGCCAGGCTATTTGGCTCAAGTCCAGAGAAATGAGGCTGGG	590							
QY	541	GAGGAGCAGAAACCAATCTTGGCTCAAGTCCAGAGAAATGAGGCTGGG	600							
DB	591	GAGGAGCAGAAACCAATCTTGGCTCAAGTCCAGAGAAATGAGGCTGGG	650							
QY	601	TGCTTCCAGGAGCCAGGCTGGGCTCAAGTCCAGAGAAATGAGGCTGGG	660							
DB	651	TGCTTCCAGGAGCCAGGCTGGGCTCAAGTCCAGAGAAATGAGGCTGGG	710							
QY	661	AGTGAAGACAGAGACTTACATCTCAATGGCTCAAGGCTGGGCTGGG	720							
DB	711	AGTGAAGACAGAGACTTACATCTCAATGGCTCAAGGCTGGGCTGGG	770							
QY	721	CTGGCAATATTTTACTGTGTTTCCAAAGACTGAGGCTGGGCTGGG	780							
DB	771	CTGGCAATATTTTACTGTGTTTCCAAAGACTGAGGCTGGGCTGGG	830							
QY	781	AAAGACAAATCACAGCATTCATAGTGAAGAGACTTGGTGGGCTGGG	840							
DB	831	AAAGACAAATCACAGCATTCATAGTGAAGAGACTTGGTGGGCTGGG	890							
QY	841	CCGGAAGATTAATAGGCTTGGGCTTCCCAACACTTGTGAAGTCCATTTTGA	900							
DB	891	CCGGAAGATTAATAGGCTTGGGCTTCCCAACACTTGTGAAGTCCATTTTGA	950							
QY	901	AAGATACCTGTGGAACATCTTGGAGAGCTGGGCTTGAAGTGGGCTTGA	960							
DB	951	AAGATACCTGTGGAACATCTTGGAGAGCTGGGCTTGAAGTGGGCTTGA	1010							

QY	961	ATCTCAACAGCGCGGTTCAGCATGGCGAGCGTCTGGCTGGCTGCTCAAGAGATTG	1020							
DB	1011	ATCTCAACAGCGCGGTTCAGCATGGCGAGCGTCTGGCTGGCTGCTCAAGAGATTG	1070							
QY	1021	ATTGAATGACTGCTGAGTACGCTGCACAAGAAACAGTTTAAACAAGAGCTCAAGTGA	1080							
DB	1071	ATTGAATGACTGCTGAGTACGCTGCACAAGAAACAGTTTAAACAAGAGCTCAAGTGA	1130							
QY	1081	TTTGGATTTGATTTCAGGAGAAATTTGCACCTGATGGCTCAGAACAGCTTACGTATG	1140							
DB	1131	TTTGGATTTGATTTCAGGAGAAATTTGCACCTGATGGCTCAGAACAGCTTACGTATG	1190							
QY	1141	ATGACCTACCTCACAGCAGGATGCTGACCAACCTGCTTCCCGACTGCTCCATCCAG	1200							
DB	1191	ATGACCTACCTCACAGCAGGATGCTGACCAACCTGCTTCCCGACTGCTCCATCCAG	1250							
QY	1201	GCAGCATGCTGAAGGTGTTTCACTCCGAGCGGCTGGCAGTGTGAGTGAGGCGCTG	1260							
DB	1251	GCAGCATGCTGAAGGTGTTTCACTCCGAGCGGCTGGCAGTGTGAGTGAGGCGCTG	1310							
QY	1261	CAGATCCTCGGGGCTTGGCTTACACAAGGACTATCCGTACGAGCGGCTACTCGGCTG	1320							
DB	1311	CAGATCCTCGGGGCTTGGCTTACACAAGGACTATCCGTACGAGCGGCTACTCGGCTG	1370							
QY	1321	ACCCGCATCCTCCTCATCTTCAGGGAACCAATGAGATTCTCCGATGTACATCGCCTG	1380							
DB	1371	ACCCGCATCCTCCTCATCTTCAGGGAACCAATGAGATTCTCCGATGTACATCGCCTG	1430							
QY	1381	ACGGCTCTGACGATGCGGCGGCTTGGCTTACACAAGGACTATCCGTACGAGCGGCT	1440							
DB	1431	ACGGCTCTGACGATGCGGCGGCTTGGCTTACACAAGGACTATCCGTACGAGCGGCT	1490							
QY	1441	AAAGTGACACAGTCATGATACGTTGGCGGAGGCTTGGGACTCCCTGGGCGGAACT	1500							
DB	1491	AAAGTGACACAGTCATGATACGTTGGCGGAGGCTTGGGACTCCCTGGGCGGAACT	1550							
QY	1501	GTGACCTGGGCTGACAGCAACCATGGAGTTGTGACCCAGCTTTGGGAGAGTGC	1560							
DB	1551	GTGACCTGGGCTGACAGCAACCATGGAGTTGTGACCCAGCTTTGGGAGAGTGC	1610							
QY	1561	AACAAGTTGAGGAGAACACCTACTGCTTGGCGGAGGCTTGGGAGAGTGC	1620							
DB	1611	AACAAGTTGAGGAGAACACCTACTGCTTGGCGGAGGCTTGGGAGAGTGC	1670							
QY	1621	TTTGGCAGACCATCATGAGGAGGAGCTGGTACTGAAGCGGTGGCCAACTCTCATC	1680							
DB	1671	TTTGGCAGACCATCATGAGGAGGAGCTGGTACTGAAGCGGTGGCCAACTCTCATC	1730							
QY	1681	AACCTGTATGGCATGACGCGCTGCTGCGGCGGAGGCTTCCATCCGATTTGGGCTC	1740							
DB	1731	AACCTGTATGGCATGACGCGCTGCTGCGGCGGAGGCTTCCATCCGATTTGGGCTC	1790							
QY	1741	CGCAACACGACGAGGTTCTTCTTGGCCAACTCTTGGCTGGGAGCTTACTTGCAG	1800							
DB	1791	CGCAACACGACGAGGTTCTTCTTGGCCAACTCTTGGCTGGGAGCTTACTTGCAG	1850							
QY	1801	AATCTCTTACGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTTAGTGAAGAT	1860							
DB	1851	AATCTCTTACGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTTAGTGAAGAT	1910							
QY	1861	AAGAAAGTGTCCAGCAGATCTCTTGGCCAACTCTTGGCTGGGAGCTTACTTGCAG	1920							
DB	1911	AAGAAAGTGTCCAGCAGATCTCTTGGCCAACTCTTGGCTGGGAGCTTACTTGCAG	1970							
QY	1921	AGGACATGCTGAGGAGGAGGAGTGTCCCTGCTACCGCGCGGCTTACCATGGCGCG	1980							
DB	1971	AGGACATGCTGAGGAGGAGGAGTGTCCCTGCTACCGCGCGGCTTACCATGGCGCG	2030							
QY	1981	TTGCTGGATGACTGTTACTCTTTTTCAGAAAGCTTGGGATTTATCAGAGTTAAGCCTT	2040							
DB	2031	TTGCTGGATGACTGTTACTCTTTTTCAGAAAGCTTGGGATTTATCAGAGTTAAGCCTT	2090							

Qy	2041	TTGTTCCCTCTGCACTGAAGGTTGCGCTGCGCTGGAGAGCCTCTCCAGGTTT	2100	CC	(b) monitoring the progression of prostate cancer in a patient;
Db	2091	TTGTTCCCTCTGCACTGAAGGTTGCGCTGCGCTGGAGAGCCTCTCCAGGTTT	2150	CC	(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
Qy	2101	TGACCTCAGCGAGTGTCTTAACAGGACATCACAGCTTCTGAACCTGAGCCGGAGAGA	2160	CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
Db	2151	TGACCTCAGCGAGTGTCTTAACAGGACATCACAGCTTCTGAACCTGAGCCGGAGAGA	2210	CC	(e) selecting a composition for inhibiting prostate cancer in a patient;
Qy	2161	GAGAAATGGAATGCTGACCCCTGGAATGGCGGGTATCTGTCATGAGGAGACACCAT	2220	CC	(f) assessing the prostate cell carcinogenic potential of a compound;
Db	2211	GAGAAATGGAATGCTGACCCCTGGAATGGCGGGTATCTGTCATGAGGAGACACCAT	2270	CC	(g) determining whether prostate cancer has metastasized in a patient;
Qy	2221	AGTGGAACTGGGCTTATCTGCTCCAGGGTGTAGGTGGGTGGGACCTGTGTC	2280	CC	(h) assessing the aggressiveness or indolence of prostate cancer in a patient;
Db	2271	AGTGGAACTGGGCTTATCTGCTCCAGGGTGTAGGTGGGTGGGACCTGTGTC	2330	CC	(i) is also useful as a pharmacodynamic or pharmacogenomic marker.
Qy	2281	AGGTGTGGATAGCCATTCTGCTCAACACACATCTCTTAAGAAACAGCTTGAACCTCT	2340	XX	Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;
Db	2331	AGGTGTGGATAGCCATTCTGCTCAACACACATCTCTTAAGAAACAGCTTGAACCTCT	2390	Qy	Query Match 100.0%; Score 2452; DB 23; Length 2522;
Qy	2341	GTCTGGGTCAATTTAACTAGAACGAGGACCTTAAACATGTACCAGGAACCAT	2400	Db	Best Local Similarity 100.0%; Pred. No. 0;
Db	2391	GTCTGGGTCAATTTAACTAGAACGAGGACCTTAAACATGTACCAGGAACCAT	2450	Qy	Matches 2452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	2401	TAACAAAGAAATATAAATGTCAATCTGTACTGTATAAATAAATAAATAAATAAATAA	2452	Db	1 CGTGTGTGTCTCCCTGCGCGCTAAGAGGGGAGACTGAGGCTAGGCTGGGGAACATCG
Db	2451	TAACAAAGAAATATAAATGTCAATCTGTACTGTATAAATAAATAAATAAATAAATAA	2502	Db	51 CGTGTGTGTCTCCCTGCGCGCTAAGAGGGGAGACTGAGGCTAGGCTGGGGAACATCG
RESULT 5				Qy	61 GGCAGCATGAGCGGCTGCGGGCTCTTCTGCGCACCCAGGCTGCGGCTGCGCTGCCGG
ABV29355				Db	111 GGCAGCATGAGCGGCTGCGGGCTCTTCTGCGCACCCAGGCTGCGGCTGCGCTGCCGG
XX	ABV29355 standard; cDNA; 2522 BP.			Qy	121 GGTCTGGTGTCTTACCGCGAACCGCGGCTACTGCGCACCCAGGCTGCGCTGACGAGCT
AC				Db	171 GGTCTGGTGTCTTACCGCGAACCGCGGCTACTGCGCACCCAGGCTGCGCTGACGAGCT
XX				Qy	181 TTGCGCAAGAGAGCTTTTCTAGGCAAAATCAAGAGAGAAAGAGTTCCTCCATTTCCAGAA
XX	16-SEP-2002 (first entry)			Db	231 TTGCGCAAGAGAGCTTTTCTAGGCAAAATCAAGAGAGAAAGATTTTCCATTTCCAGAA
DE	Human prostate expression marker cDNA 29346.			Qy	241 GTTAGCCAAAGTGAACCTTAATGAATCAATAGTCTTGGGACCCGTGGAATAATCTTC
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;			Db	291 GTTAGCCAAAGTGAACCTTAATGAATCAATAGTCTTGGGACCCGTGGAATAATCTTC
XX	pharmacogenomic marker; gene; ss.			Qy	301 ACTGAAGAGGTGGACCTCCGAAATAATGACAGGAAGGAGAAATCCAGATGAAAGCTTG
OS	Homo sapiens.			Db	351 ACTGAAGAGGTGGACCTCCGAAATAATGACAGGAAGGAGAAATCCAGATGAAAGCTTG
XX	WO200160860-A2.			Qy	361 GAGAAATTAAGAGCCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTTC
XX	23-AUG-2001.			Db	411 GAGAAATTAAGAGCCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTTC
XX	20-FEB-2001; 2001WO-US05171.			Qy	421 GGCTTCTCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGATGGTCCATC
XX	17-FEB-2000; 2000US-183319P.			Db	471 GGCTTCTCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGATGGTCCATC
XX	16-MAR-2000; 2000US-189862P.			Qy	481 ACTGTGACCTGGCAGCGCACCGGCTTATGGCTCAAGGGATCATCTTGGCTGGCCT
XX	25-MAY-2000; 2000US-207454P.			Db	531 ACTGTGACCTGGCAGCGCACCGGCTTATGGCTCAAGGGATCATCTTGGCTGGCCT
XX	09-JUN-2000; 2000US-211314P.			Qy	541 GAGGAGCAGAAAGCCAAATATCTTCCCTAAACTGCGTCCGGGAGGACATTTGAGGCTTC
XX	18-JUL-2000; 2000US-219007P.			Db	591 GAGGAGCAGAAAGCCAAATATCTTCCCTAAACTGCGTCCGGGAGGACATTTGAGGCTTC
XX	13-DEC-2000; 2000US-255281P.			Qy	601 TGCCTCACGGAGCCAGCCAGTGGGAGCGATGCGGCTCAATCCGGAGCAGAGCCACATA
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.			Db	651 TGCCTCACGGAGCCAGCCAGTGGGAGCGATGCGGCTCAATCCGGAGCAGAGCCACATA
XX	Schlegel R, Endege WO, Monahan JE;			Qy	720 AGTGAAGACAAAGAGCAGCTACATCTCAATGGCTCCAGGCTCGGATCTGATTAATGGGGA
XX	WPI; 2001-662795/76.			Db	770 AGTGAAGACAAAGAGCAGCTACATCTCAATGGCTCCAGGCTCGGATCTGATTAATGGGGA
XX	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -			Qy	780 CTGCCCAATATTTTACTGTTTCAAGAGCTGAGGCTGCTGATTTCTGATGATGATCAGTG
XX	Claim 1; Page 6287-6288; 11750pp; English.			Db	771 CTGCCCAATATTTTACTGTTTCAAGAGCTGAGGCTGCTGATTTCTGATGATCAGTG
XX	The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:			Qy	781 AAAGACAAAATCACAGCATTCATAGTAAAGAGAGACTTTGGTGGAGTCACTAATGGGAAA
XX	(a) assessing whether a patient is afflicted with prostate cancer;			Db	831 AAAGACAAAATCACAGCATTCATAGTAAAGAGAGACTTTGGTGGAGTCACTAATGGGAAA

```
QY 841 CCCGAAGATAAATTAGGCATTCGGGCTCCAAACACTGTGAAGTCCATTTTGAAGACACC 900
DB 891 CCCGAAGATAAATTAGGCATTCGGGCTCCAAACACTGTGAAGTCCATTTTGAAGACACC 950
QY 901 AAGATACCTGTGGAAACATCCTTTGGAGAGGTCCGAGATGGGTTTAAGGTGGCCATGAAC 960
DB 951 AAGATACCTGTGGAAACATCCTTTGGAGAGGTCCGAGATGGGTTTAAGGTGGCCATGAAC 1010
QY 961 ATCTCTAACAGCGCCGGTTCAGCATGGCGAGCGTCTGGCTGGGCTGCTCAAGAGATTG 1020
DB 1011 ATCTCTAACAGCGCCGGTTCAGCATGGCGAGCGTCTGGCTGGGCTGCTCAAGAGATTG 1070
QY 1021 ATTGAATGACTGCTGAGTACGCCCTGCACAAGGAACAGTTTAACAGAGGCTCACTGAA 1080
DB 1071 ATTGAATGACTGCTGAGTACGCCCTGCACAAGGAACAGTTTAACAGAGGCTCACTGAA 1130
QY 1081 TTGTGGATTGATTCAGGAGAAATTTGCACATGGCTCAGAGGCTTACGTCATGGAGAGT 1140
DB 1131 TTGTGGATTGATTCAGGAGAAATTTGCACATGGCTCAGAGGCTTACGTCATGGAGAGT 1190
QY 1141 ATGACCTACCTCACAGCAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAG 1200
DB 1191 ATGACCTACCTCACAGCAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAG 1250
QY 1201 GCAGCCATGCTGAAGGTGTTACGCTCCGAGGCGGCTGCGAGTGTGTGAGTGAGGCGCTG 1260
DB 1251 GCAGCCATGCTGAAGGTGTTACGCTCCGAGGCGGCTGCGAGTGTGTGAGTGAGGCGCTG 1310
QY 1261 CAGATCCTCGGGGCTTGGCTACACAAGGGACTATCCGTACGAGGCGATCTGGTGAC 1320
DB 1311 CAGATCCTCGGGGCTTGGCTACACAAGGGACTATCCGTACGAGGCGATCTGGTGAC 1370
QY 1321 ACCCGCATCTCTCATCTTCAGGGAACCAATGAGATTTCCGGATGTACATCCCGCTG 1380
DB 1371 ACCCGCATCTCTCATCTTCAGGGAACCAATGAGATTTCCGGATGTACATCCCGCTG 1430
QY 1381 ACGGCTCTCAGCATCGCGCGCATCTCTGACTACAGATCCATGAGCTTAAACAGGCC 1440
DB 1431 ACGGCTCTCAGCATCGCGCGCATCTCTGACTACAGATCCATGAGCTTAAACAGGCC 1490
QY 1441 AAGTGAGCACAGTATGATACCTGTGGCGGAGGCTTCGGGACTCCTCGGCGGAACT 1500
DB 1491 AAGTGAGCACAGTATGATACCTGTGGCGGAGGCTTCGGGACTCCTCGGCGGAACT 1550
QY 1501 GTGGACCTGGGCTCACAGGCAACCATGGAGTTGTGCACCCAGTCTTCGGGACAGTGC 1560
DB 1551 GTGGACCTGGGCTCACAGGCAACCATGGAGTTGTGCACCCAGTCTTCGGGACAGTGC 1610
QY 1561 AACAAAGTTGAGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACACTGCTGCTCGC 1620
DB 1611 AACAAAGTTGAGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACACTGCTGCTCGC 1670
QY 1621 TTTGGCAAGACCATCATGAGGAGAGTGTGTACTGAAGCGGGTGGCCAAACATCCTCATC 1680
DB 1671 TTTGGCAAGACCATCATGAGGAGAGTGTGTACTGAAGCGGGTGGCCAAACATCCTCATC 1730
QY 1681 AACCTGTATGGCATGACGGCGCTGCTGTCCGGCGGACCGCTCCATCCGATGGGCTC 1740
DB 1731 AACCTGTATGGCATGACGGCGCTGCTGTCCGGCGGACCGCTCCATCCGATGGGCTC 1790
QY 1741 CGCAACACGACACGAGGTTCTCTTGGCCAAACACCTTCTGCGTGGAGGCTTACTTGCAG 1800
DB 1791 CGCAACACGACACGAGGTTCTCTTGGCCAAACACCTTCTGCGTGGAGGCTTACTTGCAG 1850
QY 1801 AATCTCTTACGCTCTCTCAGCTGGACAAGTATGCTCCAGAAACCTAGATGAGCAGATT 1860
DB 1851 AATCTCTTACGCTCTCTCAGCTGGACAAGTATGCTCCAGAAACCTAGATGAGCAGATT 1910
QY 1861 AAGAAAGTGTCCACGAGATCTCTTGAAGAGCGGCTTATATCTGTGCCACCTCTGGAC 1920
DB 1911 AAGAAAGTGTCCACGAGATCTCTTGAAGAGCGGCTTATATCTGTGCCACCTCTGGAC 1970
```

```
QY 1921 AGGACATGCTGAGGAGGAGCAGATGTCCTCTGTACCGCGCCCTACCCATGCGCCG 1980
DB 1971 AGGACATGCTGAGGAGGAGCAGATGTCCTCTGTACCGCGCCCTACCCATGCGCCG 2030
QY 1981 TTGCTGGATGACTGTACTCTTTTTCAGAAAGGTGTTGGATTTATCACAGGTTAAAGCCTT 2040
DB 2031 TTGCTGGATGACTGTACTCTTTTTCAGAAAGGTGTTGGATTTATCACAGGTTAAAGCCTT 2090
QY 2041 TTGTTCCCGCTGTCACCTGAAGGGTGTGCGCTGGCGTGGGAGAGCCTCTTCCAGGTTT 2100
DB 2091 TTGTTCCCGCTGTCACCTGAAGGGTGTGCGCTGGCGTGGGAGAGCCTCTTCCAGGTTT 2150
QY 2101 TGACCTGAGGAGGAGTGTCTCTTAACAGAGCATTACAGCTTCTGAAGTGGAGAGAGAG 2160
DB 2151 TGACCTGAGGAGGAGTGTCTCTTAACAGAGCATTACAGCTTCTGAAGTGGAGAGAGAG 2210
QY 2161 GAGAATGAATTTGCTGACCCCTGGAAGTGGCGGTTATCTGAGTGGTCAATTTGAGGAGACCAT 2220
DB 2211 GAGAATGAATTTGCTGACCCCTGGAAGTGGCGGTTATCTGAGTGGTCAATTTGAGGAGACCAT 2270
QY 2221 AGTGAAGTGGGCTTATGCTGCTGCTCAGGCTGTGAGTGGTGGGAGCCTGTGTC 2280
DB 2271 AGTGAAGTGGGCTTATGCTGCTGCTCAGGCTGTGAGTGGTGGGAGCCTGTGTC 2330
QY 2281 AGGTGTGGATAGCATTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTCT 2340
DB 2331 AGGTGTGGATAGCATTCTGCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTCT 2390
QY 2341 GTCTGGGTCAATTCATTTAAACTAGAACAGAGGACCTTAAACATGTACCAGAACCAT 2400
DB 2391 GTCTGGGTCAATTCATTTAAACTAGAACAGAGGACCTTAAACATGTACCAGAACCAT 2450
QY 2401 TAACAAAGATATATAAATGTCAATCTGTACTGTATAAAAAA 2452
DB 2451 TAACAAAGATATATAAATGTCAATCTGTACTGTATAAAAAA 2502

RESULT 6
AAS40878
ID AAS40878 standard; cdna; 2492 BP.
XX
AC AAS40878;
XX
DT 17-DEC-2001 (first entry)
XX
DE cdna encoding novel human enzyme polypeptide #94.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01239.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
```

PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-465566/50.
XX P-PSDB; AAU23008.
DR
XX Novel polypeptides and polynucleotides useful for diagnosing,
PT preventing, treating neural, immune system, muscular, reproductive,
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT diseases
XX
XX Claim 4; SEQ ID NO 104; 1180pp; English.
PS
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of

CC	disorders including hyperproliferative disorders (e.g. cancer),
CC	immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC	(e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC	metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC	(e.g. asthma), cardiovascular disorders (e.g. atherosclerosis);
CC	blood-related disorders (e.g. haemophilia), reproductive disorders
CC	(e.g. infertility) and infectious disorders (e.g. influenza). The
CC	polynucleotides of the invention can also be used in gene therapy.
CC	AAS40785-AAS41684 represent cDNA sequences encoding for the novel human
CC	enzyme polypeptides of the invention.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences .
XX	
SQ	Sequence 2492 BP; 617 A; 621 C; 693 G; 558 T; 3 other;
	Query Match 99.9%; Score 2450.6; DB 22; Length 2492;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2450; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	2 GGTGTGTCTCCCTCGCGGCTAAGAGGGGAGACTGAGGTGAGGCTGGGAACATCGG 61
DB	
DB	2 GGTGTGTCTCCCTCGCGGCTAAGAGGGGAGACTGAGGTGAGGCTGGGAACATCGG 61
QY	62 GCAGCATGAGCGGCTGCGGGGCTTCTTCCTGCGCACACGCGTCTGCTTGCCTGCCGGG 121
DB	
DB	62 GCAGCATGAGCGGCTGCGGGGCTTCTTCCTGCGCACACGCGTCTGCTTGCCTGCCGGG 121
QY	122 GTCTGGTGTCTTACCGCGAACCGGGGCTACTTCGCGACACGCGCCGCTGTACGAGCTT 181
DB	
DB	122 GTCTGGTGTCTTACCGCGAACCGGGGCTACTTCGCGACACGCGCCGCTGTACGAGCTT 181
QY	182 TCGCCAAAGAGCTTTTCTAGCGAAATCAAGAGAAAGAGTTTCCCATTTCCAGAAG 241
DB	
DB	182 TCGCCAAAGAGCTTTTCTAGCGAAATCAAGAGAAAGAGTTTCCCATTTCCAGAAG 241
QY	242 TTAGCCAGATGAATTAAGTAATCAATCAGTTCTTTGGGACCGGTGGAAAAATTTCTTCA 301
DB	
DB	242 TTAGCCAGATGAATTAAGTAATCAATCAGTTCTTTGGGACCGGTGGAAAAATTTCTTCA 301
QY	302 CTGAAGAGGTGGACTCCCGAAAAATTCACAGGAAGGAAAAATCCAGATGAACATTGG 361
DB	
DB	302 CTGAAGAGGTGGACTCCCGAAAAATTCACAGGAAGGAAAAATCCAGATGAACATTGG 361
QY	362 AGAAATTGAAGAGCCTAGGGCTTTTGGGTGCAAGTCCAGAGAATATGGTGGCCTGG 421
DB	
DB	362 AGAAATTGAAGAGCCTAGGGCTTTTGGGTGCAAGTCCAGAGAATATGGTGGCCTGG 421
QY	422 GTTCTCCACACCATCTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCA 481
DB	
DB	422 GTTCTCCACACCATCTACTCAGACTAGGGGAGATCATCAGCATGGATGGTCCATCA 481
QY	482 CTGTGACCCCTGGCAGCGCACCGGCTATTGGCCTCAAGGGGATCATCTTGGTGGCATTG 541
DB	
DB	482 CTGTGACCCCTGGCAGCGCACCGGCTATTGGCCTCAAGGGGATCATCTTGGTGGCATTG 541
QY	542 AGGAGCAGAAAGCCAAATCTTGCCTAAACTGCGCTCCGGGGAGCATTTGAGCCTTCT 601
DB	
DB	542 AGGAGCAGAAAGCCAAATCTTGCCTAAACTGCGCTCCGGGGAGCATTTGAGCCTTCT 601
QY	602 GCCTCAGGAGCGAGCCAGTGGGAGCGATGCGACGCTCAATCCGGAGCAGACCACTAA 661
DB	
DB	602 GCCTCAGGAGCGAGCCAGTGGGAGCGATGCGACGCTCAATCCGGAGCAGACCACTAA 661
QY	662 GTGAAGACAAAGAGCACTACATCTTCAATGGCTCCAAAGTCTGGATTACTAATGAGGAC 721
DB	
DB	662 GTGAAGACAAAGAGCACTACATCTTCAATGGCTCCAAAGTCTGGATTACTAATGAGGAC 721
QY	722 TGCCCAATATTTTACTGTGTTTGAAGAGCACTGAGGTGCTGATTCTGATGGATCAGTGA 781
DB	
DB	722 TGCCCAATATTTTACTGTGTTTGAAGAGCACTGAGGTGCTGATTCTGATGGATCAGTGA 781
QY	782 AAGACAAAATCAGACATTCATAGTAAAGAGACTTTGGTGGAGTCTACTAATGGGAAC 841

Db	181	AAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGATTTCCTCCATTTCCAGAAGTTAGCC	240
QY	248	AAGATGAACCTTAATGAATCAATCAGTTCTTTGGGACCCGTTGGAGAAATTTCTTCACAG	307
Db	241	AAGATGAACCTTAATGAATCAATCAGTTCTTTGGGACCCGTTGGAGAAATTTCTTCACAG	300
QY	308	AGGTGGACTCCCGAAAAATTGACCAAGGAGGAAATCCCAAGATCAAACTTTTGAGAAAT	367
Db	301	AGGTGGACTCCCGAAAAATTGACCAAGGAGGAAATCCCAAGATCAAACTTTTGAGAAAT	360
QY	368	TGAACAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGGCTTCT	427
Db	361	TGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGGCTTCT	420
QY	428	CCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACTGTGA	487
Db	421	CCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACTGTGA	480
QY	488	CCCTGGCAGCCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGC	547
Db	481	CCCTGGCAGCCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGC	540
QY	548	AGAAAGCCAAATACTTGCTTAACTGGCTCCGGGAGCAGATTCGACACTTCTGCCTCA	607
Db	541	AGAAAGCCAAATACTTGCTTAACTGGCTCCGGGAGCAGATTCGACACTTCTGCCTCA	600
QY	608	CGGAGCCAGCCAGTGGGAGGATGACGCCCTCAATCCGGAGCAGAGCCACACTAAAGTGAAG	667
Db	601	CGGAGCCAGCCAGTGGGAGGATGACGCCCTCAATCCGGAGCAGAGCCACACTAAAGTGAAG	660
QY	668	ACAAGAAGCACTACATCCTCAANTGGCTCCAGGTCGAGTTACTTAATGGAGGACTGGCCA	727
Db	661	ACAAGAAGCACTACATCCTCAANTGGCTCCAGGTCGAGTTACTTAATGGAGGACTGGCCA	720
QY	728	ATAATTTTACTGTGTTGCAAGACTGAGGTGCTTGATTTCTGATGATCAGTCAAGAGCA	787
Db	721	ATAATTTTACTGTGTTGCAAGACTGAGGTGCTTGATTTCTGATGATCAGTCAAGAGCA	780
QY	788	AAATCAGACATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAAG	847
Db	781	AAATCAGACATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAAG	840
QY	848	ATAAATAGGCATTCGGGGCTCCACACACTTGTGAAGTCCATTTTGAACACACAGATAC	907
Db	841	ATAAATAGGCATTCGGGGCTCCACACACTTGTGAAGTCCATTTTGAACACACAGATAC	900
QY	908	CTGTGGAACATCCTTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATCCTCA	967
Db	901	CTGTGGAACATCCTTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATCCTCA	960
QY	968	ACAGCGCGGTTTACGATGGGAGGCTGCTGGCTGGGCTGCTCAAGAGATTGATGAAA	1027
Db	961	ACAGCGCGGTTTACGATGGGAGGCTGCTGGCTGGGCTGCTCAAGAGATTGATGAAA	1020
QY	1028	TGACTGCTGAGTACGCTGCACAAGGAAACAGTTTAAAGAGGCTTACGTCATGAGAGTATGACCT	1087
Db	1021	TGACTGCTGAGTACGCTGCACAAGGAAACAGTTTAAAGAGGCTTACGTCATGAGAGTATGACCT	1080
QY	1088	TGATTACAGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTCATGAGAGTATGACCT	1147
Db	1081	TGATTACAGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTCATGAGAGTATGACCT	1140
QY	1148	ACCTCACACAGGATGCTGGACCAACCTGGCTTTCCGACATGCTCCATCGAGGAGGCA	1207
Db	1141	ACCTCACACAGGATGCTGGACCAACCTGGCTTTCCGACATGCTCCATCGAGGAGGCA	1200
QY	1208	TGGTGAAGTGTTCAGCTCCGAGGCGGCTGGCAGTGTGAGTGAGGCGCTGAGATCC	1267
Db	1201	TGGTGAAGTGTTCAGCTCCGAGGCGGCTGGCAGTGTGAGTGAGGCGCTGAGATCC	1260
QY	1268	TCGGGGGCTTGGGCTACACAGAGGACTATCCGTACGAGCGCATACTGGGTGACACCCGCA	1327
Db	1261	TCGGGGGCTTGGGCTACACAGAGGACTATCCGTACGAGCGCATACTGGGTGACACCCGCA	1320
QY	1328	TCCTCTCATCTTTCGAGGGAACCAATCAGATTCTCCGGATGTACATCGCCCTGACGGGTC	1387
Db	1321	TCCTCTCATCTTTCGAGGGAACCAATCAGATTCTCCGGATGTACATCGCCCTGACGGGTC	1380
QY	1388	TGACGATCGCGGCCCATCTCTGACTACCAAGATTCATGAGCTTAAACAGGCAAAAGTGA	1447
Db	1381	TGACGATCGCGGCCCATCTCTGACTACCAAGATTCATGAGCTTAAACAGGCAAAAGTGA	1440
QY	1448	GCACACTCATGGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGACC	1507
Db	1441	GCACACTCATGGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGACC	1500
QY	1508	TGGGGTGCAGGCAACCATGGAGTTGTACCCAGTCTTCCGAGACAGTGCACAAGT	1567
Db	1501	TGGGGTGCAGGCAACCATGGAGTTGTACCCAGTCTTCCGAGACAGTGCACAAGT	1560
QY	1568	TTGAGGAGAACACTTACTGCTTCGGCGGACCGCTGGAGACACTGCTGCTCGCTTTGGCA	1627
Db	1561	TTGAGGAGAACACTTACTGCTTCGGCGGACCGCTGGAGACACTGCTGCTCGCTTTGGCA	1620
QY	1628	AGACCATCATGAGGAGGAGCTGCTGACTGAAGCGGTTGGCCAAACATCTCATCAACCTGT	1687
Db	1621	AGACCATCATGAGGAGGAGCTGCTGACTGAAGCGGTTGGCCAAACATCTCATCAACCTGT	1680
QY	1688	ATGGCATGACGGCCCTGCTGTCGGCGGACCGCTCCATCCGATTTGGGCTCGCAAC	1747
Db	1681	ATGGCATGACGGCCCTGCTGTCGGCGGACCGCTCCATCCGATTTGGGCTCGCAAC	1740
QY	1748	ACGACACAGAGTTCTTCTTTGGCCAAACACTTCTGCGTGGAAAGCTTACTTTCGAGAACTCT	1807
Db	1741	ACGACACAGAGTTCTTCTTTGGCCAAACACTTCTGCGTGGAAAGCTTACTTTCGAGAACTCT	1800
QY	1808	TCAGCCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATTAAAGAAAG	1867
Db	1801	TCAGCCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATTAAAGAAAG	1860
QY	1868	TGTCCTCCACAGACTCTTTCAGAGGAGGCTATATCTGTGCCACCCCTCTTGACAGGACAT	1927
Db	1861	TGTCCTCCACAGACTCTTTCAGAGGAGGCTATATCTGTGCCACCCCTCTTGACAGGACAT	1920
QY	1928	GCTGAGGAGGAGGAGTGTCCCTGCTTACCGCCGCGCCCTTACCCTGCGGCGGCTGCTGG	1987
Db	1921	GCTGAGGAGGAGGAGTGTCCCTGCTTACCGCCGCGCCCTTACCCTGCGGCGGCTGCTGG	1980
QY	1988	ATGACTGTTACTCTTTTTCAGAAAGGTTGGGATTTACAGAGTTAAGCCTTTTGTTC	2047
Db	1981	ATGACTGTTACTCTTTTTCAGAAAGGTTGGGATTTACAGAGTTAAGCCTTTTGTTC	2040
QY	2048	CCGTCTGCACCTGAAGGTTGTCCTCCCTGGCTGGGAGGCTCTTCCAGGTTTTCAGCTG	2107
Db	2041	CCGTCTGCACCTGAAGGTTGTCCTCCCTGGCTGGGAGGCTCTTCCAGGTTTTCAGCTG	2100
QY	2108	CAGCAGTGTCTCTAAACAGGACCATCACAGCTTCTGAAGTGGGAGGAGAGAGAGATG	2167
Db	2101	CAGCAGTGTCTCTAAACAGGACCATCACAGCTTCTGAAGTGGGAGGAGAGAGAGATG	2160
QY	2168	GAATTGCTGACCCCTGAACTGGGGGTTTCTGTTGTTATTGAGGAGACACCATAGTGGAA	2227
Db	2161	GAATTGCTGACCCCTGAACTGGGGGTTTCTGTTGTTATTGAGGAGACACCATAGTGGAA	2220
QY	2228	ACTGGGGCTTATGCTGCTGCTCCAGGCTGAGGTGGGTGGGAGCTGCTGTCAGGTTG	2287
Db	2221	ACTGGGGCTTATGCTGCTGCTCCAGGCTGAGGTGGGTGGGAGCTGCTGTCAGGTTG	2280
QY	2288	GATAGCCATTTGCTCAACACACTTCTTCTAAGAAACAGCTTGAAGCTCTGTCTGG	2347
Db	2281	GATAGCCATTTGCTCAACACACTTCTTCTAAGAAACAGCTTGAAGCTCTGTCTGG	2340
QY	2348	TCATTTCATTTAAACTAGAAGCAGGACTTAAACATGTACAGAGAACCATTTTAAACAA	2407
Db	2341	TCATTTCATTTAAACTAGAAGCAGGACTTAAACATGTACAGAGAACCATTTTAAACAA	2400

Qy	2408	GAATATAAAATGTCACAATCTGTGTA	CTGTTAAAAA	2447
Dh	2401	GAATATAAAATGTCACAATCTGTGTA <th>CTGTTAAAAA</th> <th>2440</th>	CTGTTAAAAA	2440

RESULT 8

RESOL 8
AAH16781
ID AAH16781 standard: cDNA: 2399 BP.

AA
AC
AAH16781:

26-JUN-2001 (first entry)

Human cDNA sequence SEO ID NO:16009.

xx human: primer: detection: diagnosis: antisense therapy: gene therapy; ss.

XX
50

XX
DN
FD1074617-A2XX
DD
07-FEB-2001XX
DE
38 - TTT - 2000. 2000EP-0116136

XX
00-7777-1000.
00TP-0249036

PR 27-AUG-1999; 99JP-0300253.
PR 27-AUG-1999; 99JP-0300253.
PR 27-AUG-1999; 99JP-0300253.

02-MAY-2000; 2000JP-0183767.

	E	C	F	C	B	C	Z	H	D	:	A	H	M	:	S	X	X
--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

XX

PI Ishii S, Sugiyama T,

DR WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly

PT and/or diagnosis of

[illegible]

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13029 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

sequence 2399 BP: 597 A; 606 C; 672 G; 524 T; 0 other;

Query Match 97.2%; Score 2382.2; DB 22; Length 2399;
Best Local Similarity 99.8%;
Matches 2396: Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY	21	GCTAAGAGGGAGACTGAGCTGAGCGTGGGGAACATCGGCAGCATGAGCGGCTGCGG	80
DB	1	GCTAAGAGGGAGACTGAGCGTGGGGAACATCGGCAGCATGAGCGGCTGCGG	60
QY	81	GCTCTTCTCGCACACAGCGCTGCGGCTCGTGCCTGCGGGGTCTGTGTGCTCTACCGC	140
DB	61	GCTCTTCTCGCACACAGCGCTGCGGCTCGTGCCTGCGGGGTCTGTGTGCTCTACCGC	120
QY	141	GAACCGGCGCTACTGCGCACAGCGCGCTGTACGAGCTTTGCGCAAGAGCTTTTCCT	200
DB	121	GAACCGGCGCTACTGCGCACAGCGCGCTGTACGAGCTTTGCGCAAGAGCTTTTCCT	180
QY	201	AGGCAAAATCAAGAAAGAAAGTTTCCCAFTTCCAGAGTTAGCCAAGATGAACTTAA	260
DB	181	AGGCAAAATCAAGAAAGAAAGTTTCCCAFTTCCAGAGTTAGCCAAGATGAACTTAA	240
QY	261	TGAATCAATCAGTTCTTGGGACCGTGAAAAATTTCTCACTGAAGAGTTGGAATCCCG	320
DB	241	TGAATCAATCAGTTCTTGGGACCGTGAAAAATTTCTCACTGAAGAGTTGGAATCCCG	300
QY	321	AAAAATTGACCAAGGAAGGAAATCCAGATGAAATTTGGAGAAATTTGAAGAGCTTAGG	380
DB	301	AAAAATTGACCAAGGAAGGAAATCCAGATGAAATTTGGAGAAATTTGAAGAGCTTAGG	360
QY	381	GCCTTTTGGGCTGCAAGTCCCAGAGAAATATGTGGCTGGGCTTCTCCACACCATGTA	440
DB	361	GCCTTTTGGGCTGCAAGTCCCAGAGAAATATGTGGCTGGGCTTCTCCACACCATGTA	420
QY	441	CTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCATGTGACCTTGGCAGCGCA	500
DB	421	CTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCATGTGACCTTGGCAGCGCA	480
QY	501	CCAGGCTATTGGGCTCAAGGGATCATTTGGCTGGCCTGAGGAGCAGAAAGCCAAATA	560
DB	481	CCAGGCTATTGGGCTCAAGGGATCATTTGGCTGGCCTGAGGAGCAGAAAGCCAAATA	540
QY	561	CTTGCCCTAAACTGCGCTCGGGGAGCACATTCAGGCGCTTCTGGCTCAGGAGCCAGCG	620
DB	541	CTTGCCCTAAACTGCGCTCGGGGAGCACATTCAGGCGCTTCTGGCTCAGGAGCCAGCG	600
QY	621	TGGGAGCGATCAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAAGACAAGACACTA	680
DB	601	TGGGAGCGATCAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAAGACAAGACACTA	660
QY	681	CATCCTCAATGGCTCCAAGGCTGGATTACTAATGAGGACTGGCCCAATATTTTACTGT	740
DB	661	CATCCTCAATGGCTCCAAGGCTGGATTACTAATGAGGACTGGCCCAATATTTTACTGT	720
QY	741	GTTTGCAAGACTGAGGTCGTTGATTCTGTGATGATCAGTGAAGACAAAATCACAGATT	800
DB	721	GTTTGCAAGACTGAGGTCGTTGATTCTGTGATGATCAGTGAAGACAAAATCACAGATT	780
QY	801	CATAGTAGAAGAGACTTTGGTGAGTCACTAATGGGAACCCGAGAGATAAATTAGGCAT	860
DB	781	CATAGTAGAAGAGACTTTGGTGAGTCACTAATGGGAACCCGAGAGATAAATTAGGCAT	840
QY	861	TCGGGGCTCCAACACTTGTGAAGTCCATTTTGAACACACCAAGATACCTGTGAAACAT	920
DB	841	TCGGGGCTCCAACACTTGTGAAGTCCATTTTGAACACACCAAGATACCTGTGAAACAT	900
QY	921	CCTTGGAGAGTTCGGAGATGGTTTAAGTGGCCATGAACATCCTCAACAGCGCCGGTT	980
DB	901	CCTTGGAGAGTTCGGAGATGGTTTAAGTGGCCATGAACATCCTCAACAGCGCCGGTT	960
QY	981	CAGCATGGCAGCGTCGTGGCTGGGCTGCTCAAGAGATTGATTAATGACTGCTGAGTA	1040
DB	961	CAGCATGGCAGCGTCGTGGCTGGGCTGCTCAAGAGATTGATTAATGACTGCTGAGTA	1020
QY	1041	CGCCTGCACAAGAAACAGTTTAACAAGAGGCTCAGTCAATTTTGATTTGATTCAGGAGAA	1100
DB	1021	CGCCTGCACAAGAAACAGTTTAACAAGAGGCTCAGTCAATTTTGATTTGATTCAGGAGAA	1080

QY 1101 ATTTCACATGCTCAGAGGCTTACGTATGAGAGTATGACCTACCTACAGCAGG 1160
DB 1081 ATTTCACATGCTCAGAGGCTTACGTATGAGAGTATGACCTACCTACAGCAGG 1140
QY 1161 GATGCTGGACCAACCTGGCTTCCGACTGCTCCATGAGGAGCCATGTTGAAGTGT 1220
DB 1141 GATGCTGGACCAACCTGGCTTCCGACTGCTCCATGAGGAGCCATGTTGAAGTGT 1200
QY 1221 CAGCTCCGAGGCGCCCTGGCAGTGTGAGTGAAGGCTGACATCCTCGGGGCTTGGG 1280
DB 1201 CAGCTCCGAGGCGCCCTGGCAGTGTGAGTGAAGGCTGACATCCTCGGGGCTTGGG 1260
QY 1281 CTACACAGGAGTATCCGTACGAGGCGATPACTGCTGACACCCGATCCTCCTCATCTT 1340
DB 1261 CTACACAGGAGTATCCGTACGAGGCGATPACTGCTGACACCCGATCCTCCTCATCTT 1320
QY 1341 CGAGGGAACCAATGAGATCTCCGATGTACATCGCCCTGACGGGCTGACGATCCGG 1400
DB 1321 CGAGGGAACCAATGAGATCTCCGATGTACATCGCCCTGACGGGCTGACGATCCGG 1380
QY 1401 CCGCATCCTGACTACGAGTATCCATGAGCTTAAACAGGCGCAAGTGCACACATCATGA 1460
DB 1381 CCGCATCCTGACTACGAGTATCCATGAGCTTAAACAGGCGCAAGTGCACACATCATGA 1440
QY 1461 TACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCCGAAGTGTGACCTGGGGCTGACAGG 1520
DB 1441 TACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCCGAAGTGTGACCTGGGGCTGACAGG 1500
QY 1521 CAACCATGGAGTTGTCACCCAGTCTCGGACAGTGCACCAAGTTTGAGGAGAACAC 1580
DB 1501 CAACCATGGAGTTGTCACCCAGTCTCGGACAGTGCACCAAGTTTGAGGAGAACAC 1560
QY 1581 CTACTGCTTCGGCCGAGGCTGGAGACTGCTGCTCCGCTTTGGCAAGACCATCATGA 1640
DB 1561 CTACTGCTTCGGCCGAGGCTGGAGACTGCTGCTCCGCTTTGGCAAGACCATCATGA 1620
QY 1641 GGAGCAGCTGGTACTGAACGGGGTGCACCAATCTCATCAACCTGTATGGCATACGGC 1700
DB 1621 GGAGCAGCTGGTACTGAACGGGGTGCACCAATCTCATCAACCTGTATGGCATACGGC 1680
QY 1701 CGTGTGTCGGGGCAGCGCTCCATCGGATGGGCTCCGCAACCGACGACGAGGT 1760
DB 1681 CGTGTGTCGGGGCAGCGCTCCATCGGATGGGCTCCGCAACCGACGACGAGGT 1740
QY 1761 TCTCTTGGCCACACCTTCGGTGGAGCTTACTTGCAGAACTCTCTCAGCCTCTCTCA 1820
DB 1741 TCTCTTGGCCACACCTTCGGTGGAGCTTACTTGCAGAACTCTCTCAGCCTCTCTCA 1800
QY 1821 GCTGGACAAGTATGCTCCAGAAACCTTAGATGAGCAGATTAAGAAGTGTCCAGCAGAT 1880
DB 1801 GCTGGACAAGTATGCTCCAGAAACCTTAGATGAGCAGATTAAGAAGTGTCCAGCAGAT 1860
QY 1881 CTTTGAAGGAGGCTATATCTGTGCCACCCCTCTGGACAGGACATGCTGAGGAGGGG 1940
DB 1861 CTTTGAAGGAGGCTATATCTGTGCCACCCCTCTGGACAGGACATGCTGAGGAGGGG 1920
QY 1941 ACAGGTGCTCCCTGCTACCGCCCGCCCTACCCATGGCCGCTGTGGATGACTGTACTC 2000
DB 1921 ACAGGTGCTCCCTGCTACCGCCCGCCCTACCCATGGCCGCTGTGGATGACTGTACTC 1980
QY 2001 TTTTTCAGAAAGGTGTTGGGATTAATCAGAGTAAAGCCTTTTGTTCGCCGCTGCACCTG 2060
DB 1981 TTTTTCAGAAAGGTGTTGGGATTAATCAGAGTAAAGCCTTTTGTTCGCCGCTGCACCTG 2040
QY 2061 AAGGTTGCTCGCTGGGAGGAGGCTCTTCAGGTTTTCAGCTGCGAGGAGTGTCT 2120
DB 2041 AAGGTTGCTCGCTGGGAGGAGGCTCTTCAGGTTTTCAGCTGCGAGGAGTGTCT 2100
QY 2121 CTACAGGACCATCACAGCTTCTCACTGAGCCCGGAGAGAGAGATGGAATTGCTGACCC 2180
DB 2101 CTACAGGACCATCACAGCTTCTCACTGAGCCCGGAGAGAGATGGAATTGCTGACCC 2160
QY 2181 CTGGAAGTGGCGGCTATCTGTGCTATGAGGAGACACATAGTGAAGCTGGGCTTATG 2240

DB 2159 CTGGAAGTGGCGGCTATCTGTGCTATGAGGAGACACCATAGTGAAGCTGGGCTTATG 2218
QY 2241 CTGCTCCCTCCAGGCTGAGGTGGGTGGGACCTGTGTCAGGTGTGGATAGCCATTCT 2300
DB 2219 CTGCTCCCTCCAGGCTGAGGTGGGTGGGACCTGTGTCAGGTGTGGATAGCCATTCT 2278
QY 2301 GCTCAACACACATCTCTTAAGAAACAGCTTGAAGAGCTGTGCTGGGTCAATCAATTA 2360
DB 2279 GCTCAACACACATCTCTTAAGAAACAGCTTGAAGAGCTGTGCTGGGTCAATCAATTA 2338
QY 2361 CTAGAGCAGAGGCTTAAACATGTACCAGGACCATTTAAACAAAGATATAAATGT 2420
DB 2339 CTAGAGCAGAGGCTTAAACATGTACCAGGACCATTTAAACAAAGATATAAATGT 2398
QY 2421 C 2421
DB 2399 C 2399

RESULT 9
AAC76009

ID AAC76009 standard; cdna: 2218 BP.

XX AAC76009;

XX AC AAC76009;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF1564 polynucleotide sequence SEQ ID NO:3127.

XX KW Human: open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX KW vulnary; antipariatic; antiparkinsonian; nootropic; neuroprotective;
XX KW anticonvulsant; osteopathic; antilithritic; immunosuppressant; cardiant;
XX KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
XX KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX KW cholesterol ester storage; systemic lupus erythematosus; infection;
XX KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX KW thrombosis; contraceptive; ss.

XX OS Homo sapiens.

XX PN WO200058473-A2.

XX PD 05-OCT-2000.

XX PF 31-MAR-2000; 2000WO-US08621.

XX PR 31-MAR-1999; 99US-0127607.

XX PR 02-APR-1999; 99US-0127636.

XX PR 30-MAR-2000; 2000US-0540763.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX DR WPI; 2000-602362/57.

XX DR P-PSDB; AAB41800.

XX PT Novel nucleic acids and peptides derived from open reading frame X,
XX PT useful for treating e.g. cancers, proliferative disorders,
XX PS neurodegenerative disorders and cardiovascular disease -

XX CC Claim 5; Page 2343-2344; 5507pp; English.

XX CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX CC which represent the human ORFX open reading frames 1 to 3161. The ORFX

sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ Sequence 2218 BP; 573 A; 548 C; 609 G; 487 T; 1 other;

Query Match 90.3%; Score 2213.8; DB:21; Length 2218;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2215; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 236 CAGAAGTTAGCAAGATGAATTAATGAATCAATCAGTTCTTTGGACCCGCGTGAATAAT 295
Db 2 CAGAAAGTTAGCAAGATGAATTAATGAATCAATCAGTTCTTTGGACCCGCGTGAATAAT 61

Qy 296 TCTTCACTGAAGAGTGGACTCCGAAAATTTGACAGGAAGGAAAATCCAGATGAAA 355
Db 62 TCTTCACTGAAGAGTGGACTCCGAAAATTTGACAGGAAGGAAAATCCAGATGAAA 121

Qy 356 CTTTGGAGAATTAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAATATGGTG 415
Db 122 CTTTGGAGAATTAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAATATGGTG 181

Qy 416 GCCTGGGCTTCTCAACACCATTGCTCAAGACTAGGGAGATCATCAGATGGATGGGT 475
Db 182 GCCTGGGCTTCTCAACACCATTGCTCAAGACTAGGGAGATCATCAGATGGATGGGT 241

Qy 476 CCATCACTGTGACCTCGCAGCGCACCGGCTATTGGCTCAAGGGGATCATCTTGCTG 535
Db 242 CCATCACTGTGACCTCGCAGCGCACCGGCTATTGGCTCAAGGGGATCATCTTGCTG 301

Qy 536 GCATGAGGAGCAAGAACCCAAATPACTTGCTTAACATGGGCTCGGGAGCACATTGCAG 595
Db 302 GCATGAGGAGCAAGAACCCAAATPACTTGCTTAACATGGGCTCGGGAGCACATAGCAG 361

Qy 596 CCTTCTGCTCAGCGGCCAGCCAGTGGAGCGATGAGCTCAATCCGGAGCAGGCCA 655
Db 362 CCTTCTGCTCAGCGGCCAGCCAGTGGAGCGATGAGCTCAATCCGGAGCAGGCCA 421

Qy 656 CACTAAGTGAAGCAAGACACTACATCTCTCAATGGCTCCAAGGCTGGATTACTAATG 715
Db 422 CACTAAGTGAAGCAAGACACTACATCTCTCAATGGCTCCAAGGCTGGATTACTAATG 481

Qy 716 GAGGACTGGCCAAATATTTTACTGTGTGTTGCAAGACTGAGGCTGTTGATTCGATGGAT 775
Db 482 GAGGACTGGCCAAATATTTTACTGTGTGTTGCAAGACTGAGGCTGTTGATTCGATGGAT 541

Qy 776 CAGTCAAGACAAATACAGCATTCATAGTAGAAGAGACTTTTGGTGGAGTCACTAATG 835
Db 542 CAGTCAAGACAAATACAGCATTCATAGTAGAAGAGACTTTTGGTGGAGTCACTAATG 601

Qy 836 GGAACCCGAGATAAATTAGGCATTCGGGCTCCAACTGTGAAGTCAATTTTGA 895
Db 602 GGAACCCGAGATAAATTAGGCATTCGGGCTCCAACTGTGAAGTCAATTTTGA 661

Qy 896 ACACCAAGTACCTGTGGAACACATCTTGGAGAGTGGGTTTAAAGTGGCCA 955
Db 662 ACACCAAGTACCTGTGGAACACATCTTGGAGAGTGGGTTTAAAGTGGCCA 721

Qy 956 TGAACATCCTCAACAGCGCGGTTTACATATGGGAGCGTCTGGCTGGGCTGCTCAAGA 1015
Db 722 TGAACATCCTCAACAGCGCGGTTTACATATGGGAGCGTCTGGCTGGGCTGCTCAAGA 781

Qy 1016 GATTGATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAAGAGGCTCA 1075
Db 782 GATTGATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAAGAGGCTCA 841

Qy 1076 GTGAATTTGGATTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAGCTTTACGTCATGG 1135
Db 842 GTGAATTTGGATTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAAGCTTTACGTCATGG 901

Qy 1136 AGAGTATGACCTPACCTCACAGCAGGATGTGGACCAACCTGGCTTTCCGACTGCTCCA 1195
Db 902 AGAGTATGACCTPACCTCACAGCAGGATGTGGACCAACCTGGCTTTCCGACTGCTCCA 961

Qy 1196 TCGAGGCGCCATGTGAAGGTGTTTCACTCCGAGCGCCCTGGCAGTGTGTGAGTGAAG 1255
Db 962 TCGAGGCGCCATGTGAAGGTGTTTCACTCCGAGCGCCCTGGCAGTGTGTGAGTGAAG 1021

Qy 1256 CGCTGCAGATCCTCGGGGCTTTGGGCTACACAAGGAGTATCCGTACGAGCGCATCTGCG 1315
Db 1022 CGCTGCAGATCCTCGGGGCTTTGGGCTACACAAGGAGTATCCGTACGAGCGCATCTGCG 1081

Qy 1316 GTGACACCGCATCCTCCTCATCTTCGAGGGAAACCAATGAGATTTCCGGATGTACATCG 1375
Db 1082 GTGACACCGCATCCTCCTCATCTTCGAGGGAAACCAATGAGATTTCCGGATGTACATCG 1141

Qy 1376 CCTGTGACGGTCTGCAGCATGCGCGCGCATCTGACTACCGAGATPCCATGAGCTTAAAC 1435
Db 1142 CCTGTGACGGTCTGCAGCATGCGCGCGCATCTGACTACCGAGATPCCATGAGCTTAAAC 1201

Qy 1436 AGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCG 1495
Db 1202 AGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCG 1261

Qy 1496 GAACTGTGACCTGGGCTGACAGCAACCATGGAGTTGTGACCCCGCTTTGCGGAGCA 1555
Db 1262 GAACTGTGACCTGGGCTGACAGCAACCATGGAGTTGTGACCCCGCTTTGCGGAGCA 1321

Qy 1556 GTGCCAAAGTTTGAAGAGACACCTACTGTTGGCGCGGACCTGCGGAGACACTGCTGC 1615
Db 1322 GTGCCAAAGTTTGAAGAGACACCTACTGTTGGCGCGGACCTGCGGAGACACTGCTGC 1381

Qy 1616 TCCGCTTTGGCAAGACCATCATCGAGGACAGCTGCTGTAAGCGGGTGGCCAAACATCC 1675
Db 1382 TCCGCTTTGGCAAGACCATCATCGAGGACAGCTGCTGTAAGCGGGTGGCCAAACATCC 1441

Qy 1676 TCATCAACCTGTATGGCATGACGGCGTGTGTCGGGGCGCAGCCCTCCATCCGATTCG 1735
Db 1442 TCATCAACCTGTATGGCATGACGGCGTGTGTCGGGGCGCAGCCCTCCATCCGATTCG 1501

Qy 1736 GGCTCCGACACGACCGACCGAGGTTCTCTTGGCCAAACACTTCTCGCTGGAAGCTTACT 1795
Db 1502 GGCTCCGACACCGACCGACCGAGGTTCTCTTGGCCAAACACTTCTCGCTGGAAGCTTACT 1561

Qy 1796 TGCAGAACTCTTTCAGCCTCTCTCAGCTGACAAAGTATGCTCAGAAAAACCTAGATGAGC 1855
Db 1562 TGCAGAACTCTTTCAGCCTCTCTCAGCTGACAAAGTATGCTCAGAAAAACCTAGATGAGC 1621

Qy 1856 AGATTAAAGAAAGTGTCCACAGCATCTTCAGAAAGGAGGCTATATCTGTGCCACCCCTC 1915
Db 1622 AGATTAAAGAAAGTGTCCACAGCATCTTCAGAAAGGAGGCTATATCTGTGCCACCCCTC 1681

Qy 1916 TGCACAGGACATGCTGAGGCGAGGAGCAGTGTCCCTGCTTACCGCGCCGCCCCCTACCCATG 1975
Db 1682 TGCACAGGACATGCTGAGGCGAGGAGCAGTGTCCCTGCTTACCGCGCCGCCCCCTACCCATG 1741

Qy 1976 GCCGTTGCTGGATGACTGTACTCTTTTTCAGAAAGGCTTGGGATTTATCAGAGTTAA 2035
Db 1742 GCCGTTGCTGGATGACTGTACTCTTTTTCAGAAAGGCTTGGGATTTATCAGAGTTAA 1801

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0215647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465566/50.
P-PSDB; AAU23012.
Novel polypeptides and polynucleotides useful for diagnosing,
preventing, treating neural, immune system, muscular, reproductive,
pulmonary, cardiovascular, renal, proliferative disorders and cancerous
diseases
Claim 4; SEQ ID NO 108; 1180pp; English.

comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AA892446 to AA893893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the amplification of the present invention.

SQ Sequence 771 BP; 204 A; 187 C; 219 G; 158 T; 3 other;

Query Match	30.6%	Score 750.6	DB 22	Length 771
Best Local Similarity	99.0%	Pred. No. 8	2e-192	
Matches 764	Conservative 0	Mismatches 7	Indels 1	Gaps 1
Qy	21	GCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGCGACGATGACGCGCTCGCGG	80	
Db				
Qy	1	GCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGCGACGATGACGCGCTCGCGG	60	
Db				
Qy	81	GCTCTTCTCGCACACAGGGCTGCGGGCTGCTGCTGCCGGGTCTGGTGGTCTCTACCGC	140	
Db	61	GCTCTTCTCGCACACAGGGCTGCGGGCTGCTGCTGCCGGGTCTGGTGGTCTCTACCGC	120	
Qy	141	GAACCGGCGGCTACTGCGCACACAGCCCGCCTGTACGAGCTTTCGCCAAAGAGCTTTTCTC	200	
Db	121	GAACCGGCGGCTACTGCGCACACAGCCCGCCTGTACGAGCTTTCGCCAAAGAGCTTTTCTC	180	
Qy	201	AGGCAAAATCAAGAAAGAAAGTATTTTCCCATTTCCAGAAGTTAGCCAAAGATGAACTTAA	260	
Db	181	AGGCAAAATCAAGAAAGAAAGTATTTTCCCATTTCCAGAAGTTAGCCAAAGATGAACTTAA	240	
Qy	261	TCAAATCAATCAGTTCTTTGGGACCGGTGGAAAAATTTCTACTGAAGAGGTGGACTCCCG	320	
Db	241	TGAATCAATCAGTTCTTTGGGACCGGTGGAAAAATTTCTACTGAAGAGGTGGACTCCCG	300	
Qy	321	AAAAATTGACCAGGAAGGAAATCCACAGATGAACCTTTGGAGAAATTCAGAGAGCCTAGG	380	
Db	301	AAAAATTGACCAGGAAGGAAATCCACAGATGAACCTTTGGAGAAATTCAGAGAGCCTAGG	360	
Qy	381	GCTTTTGGGCTGCAAGTCCAGAAATATGTGGCCTGGGCTTCTCCAAACACCATGTA	440	
Db	361	GCTTTTGGGCTGCAAGTCCAGAAATATGTGGCCTGGGCTTCTCCAAACACCATGTA	420	
Qy	441	CTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCTGTGACCTGGCAGCGCA	500	
Db	421	CTCAGACTAGGGGAGATCATCAGCATGGATGGTCCATCTGTGACCTGGCAGCGCA	480	
Qy	501	CAAGCTATTGGCCTCAAGGGGATCATTTGGCTGGCCTGAGGAGCAGAAAGCCAAATA	560	
Db	481	CAAGCTATTGGCCTCAAGGGGATCATTTGGCTGGCCTGAGGAGCAGAAAGCCAAATA	540	
Qy	561	CTTGCTTAATCTGGCTCCGGGAGCACATTCGACGCTTTCGCTCAGGAGCCAGCGAG	620	
Db	541	CTTGCTTAATCTGGCTCCGGGAGCACATTCGACGCTTTCGCTCAGGAGCCAGCGAG	600	
Qy	621	TGGAGCGGATGCACGCTCAATCCGAGCAGAGCCACACTAAGTGAAGACAAGAGCACTA	680	
Db	601	TGGAGCGGATGCACGCTCAATCCGAGCAGAGCCACACTAAGTGAAGACAAGAGCACTA	660	
Qy	681	CATCCTCAATGGCTCCAAGGCTTGGATTACTAATGGAGGACTGGCCAAATATTTTACTGT	740	
Db				

WW/rsp5/WMP domain; therapy; forensic; genetic mapping; diagnostic; detection; treatment; cervical; melanoma; colorectal adenocarcinoma; Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma; leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal; prostate; ss.

Homo sapiens.

WO9933982-A2. 98WO-US27610.

08-JUL-1999.

22-DEC-1998; 98US-0217471.

21-DEC-1998; 97US-0068755.

23-DEC-1997; 98US-0080664.

03-APR-1998; 98US-0105234.

21-OCT-1998; 98US-0105877.

27-OCT-1998; 98US-0105877.

(CHIR) CHIRON CORP.

(HYSE-) HYSEQ INC.

Crkvenjakov R, Dickson M, Drmanac R, Drmanac S, Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA; Jones LW, Kassam A, Kennedy GC, Kita D, Labat I; Lanson G, Lesskowitz D, Pot D, Randazzo F, Reinhard C; Stache-Crain B, Sudduth-Klinger J, Williams LT; WPI; 1999-430243/36.

New isolated human polynucleotides

Claim 1; Page 503; 59lpp; English.

This invention describes novel isolated human polynucleotides obtained by screening for differential expression in colon cancer, breast cancer and lung cancer cell lines. The polynucleotides of the invention are represented in AAX9875-X99118 and encode polypeptides of protein families selected from 4 transmembrane segments integral membrane proteins, 7 transmembrane receptors, ATPases associated with various cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of transcription factors, G-protein alpha subunit, phospholipase C, diacylglycerol binding proteins, protein kinase, protein phosphatase 2C, protein tyrosine phosphatase, trypsin, wnt family of developmental signalling proteins and WW/rsp5/WMP domain containing proteins. The encoded polypeptides also have a functional domain selected from Ank repeat, basic region plus leucine zipper transcription factors, bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease domain. The polynucleotides encode polypeptides with similarity to known protein families and are predicted to have similar properties. The novel polynucleotides can be used to develop products for use as therapeutic agents and in forensics, genetic analysis, mapping and diagnostic applications. In particular, the product can be used for the detection and management of cancers. They can be used for treating e.g. cervical cancers, melanomas, colorectal adenocarcinomas, Wilm's tumour, sarcomas, retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric hereditary ectodermal dysplasia, congenital alveolar dysplasia, epithelial dysplasia of the cervix, fibrous dysplasia of bone, and mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast, prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of the skin.

Sequence 772 BP; 161 A; 207 C; 186 G; 165 T; 53 other;

Query Match 20.4%; Score 499.6; DB 20; Length 772;

Best Local Similarity 91.1%; Pred. No. 3.9e-124;

Matches 545; Conservative 0; Mismatches 49; Indels 4; Gaps 3;

Oy 1098 GAAATTTGCACTGATGGCTCAGAGGCTTACGTGATGGAGATGACCTACCTCACGC 1157

Db 83 GAAATTTGCACTGATGGCTCANAAGGCTTACGTTTGGAGAGTATGACCTACCTCACAGN 142

Qy 1158 AGGGATGCTGACCAAACTGCTTCCGACTGCTCCATCGAGGACCCATGTTGAGAGT 1217

Db 143 AGGGATGCTGACCAAACTGCTTCCGACTGCTCCATCGAGGACCCATGTTGAGAGT 202

Qy 1218 GTTCAGCTCCGAGGCGGCTGCGAGTGTGTGAGTGAGCGGCTGAGATCCTCGGGGCTT 1277

Db 203 GTTCANCTCCGAGGCGGCTGCGAGTGTGTGAGTGAGCGGCTGAGATCCTCGGGGCTN 262

Qy 1278 GGGCTACACAAGGGGACTATCCGTAGGAGCGCATCTGCTGACACCGGCTATCCTCTCAT 1337

Db 263 GGGCTACACAAGGGGACTATCCGTAGGAGCGCATCTGCTGAGCGCATCTCTTACTCAT 322

Qy 1338 CTTTCAGGGAACCAATGAGATTTCCGGATGTATCATGCCCTGAGGGTCTGAGCATGC 1397

Db 323 CTTTCNAGGGAACCAATGAGATTTCCGGATGTATCATGCCCTGAGGGTCTGAGCATGC 382

Qy 1398 CGGGCGCATCTGACTACCGAGTCCATGAGCTTAAACAGCGCAAAAGTGAGCACAGTCAT 1457

Db 383 CGGGCGCATCTGACTACCGAGTCCATGAGCTTAAACAGCGCAAAAGTGAGCACAGTCAT 442

Qy 1458 GGATACCGTTGGCGGAGGCTTTCGGGACTCCCTGGGCGCAAACTGTGGACCTGGGGCTGAC 1517

Db 443 GGATACCGTTGGCGGAGGCTTTCGGGACTTNCCTGGGCGCAAACTGTGGACCTGGGGCTGAC 502

Qy 1518 AGGCAACCATGAGTGTGACCCAGTCTTGGCGACAGTGCCAAAGTTTGGAGAGAA 1577

Db 503 AGGCAACCATGAGTGTGACCCAGTCTTGGCGACAGTGCCAAAGTTTGGAGAGAA 562

Qy 1578 CACCTACTCTTCGGC-CGGAGCTGGAGACACTGCTGCTCGGCTTTGGC-AAGACCATC 1635

Db 563 CACCTACTCTTANCTGACCGTGAGACACTTGTGNTNCCNTTTGGCAAAAGACCATC 622

Qy 1636 ATGAGGAGCAGCTGG--TACTGAGCGGGTGGCAACATCTCTCATCACTGTATGG 1691

Db 623 ATGGANGANGANNNGTGTCTNAANCNGNTGGCCAACTNCTCATCAACCTGTATTTG 680

Search completed: September 6, 2003, 17:43:01

Job time : 681.218 secs .

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: September 6, 2003, 10:14:13 ; Search time 140.926 Seconds
(without alignments)
7679.711 Million cell updates/sec

Title: US-09-945-326-1
Perfect score: 2452
Sequence: 1 cgtgtgtgtcctcctcgccg.....actgttaaaaaaaaaaaaaa 2452

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/3A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/3B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	168.6	6.9	1341	4	US-09-252-991A-11446
2	165.6	6.8	1212	4	US-09-252-991A-11389
3	155.6	6.3	1173	4	US-09-252-991A-9487
C 4	155.6	6.3	1209	4	US-09-252-991A-9346
C 5	143.4	5.8	4403765	3	US-09-103-840A-2
C 6	143.4	5.8	4411529	3	US-09-103-840A-1
C 7	139.4	5.7	3515	4	US-09-221-017B-824
C 8	134	5.5	1395	4	US-09-252-991A-16537
C 9	134	5.5	1665	4	US-09-252-991A-16061
10	132.6	5.4	33312	4	US-08-311-731A-121
11	123.2	5.0	1260	4	US-09-252-991A-14526
C 12	123.2	5.0	1260	4	US-09-252-991A-14838
13	111.2	4.5	1155	4	US-09-648-004-5
14	111.2	4.5	17417	4	US-09-648-004-27
15	105.8	4.3	1170	4	US-09-328-352-2254
16	102	4.2	1275	4	US-09-252-991A-8082
17	102	4.2	1488	4	US-09-252-991A-8036
C 18	102	4.2	2055	4	US-09-252-991A-8231
19	97	4.0	1185	4	US-09-328-352-2675
20	92.2	3.8	4403765	3	US-09-103-840A-2
21	92.2	3.8	4411529	3	US-09-103-840A-1
22	91.8	3.7	849	4	US-09-252-991A-12573
C 23	91.8	3.7	969	4	US-09-252-991A-12911
24	91.8	3.7	1257	4	US-09-252-991A-12881
25	90.4	3.7	1539	4	US-09-364-230-29
26	87.2	3.6	1164	4	US-09-328-352-2316
C 27	86.6	3.5	1440	4	US-09-252-991A-2675

28	86.6	3.5	2592	4	US-09-252-991A-3003	Sequence 3003, Ap
29	85.4	3.5	1553	4	US-09-364-230-31	Sequence 31, Appl
30	74.4	3.0	2079	4	US-09-252-991A-11426	Sequence 11426, A
31	74	3.0	1248	4	US-09-328-352-573	Sequence 573, App
32	72	2.9	1209	4	US-09-328-352-155	Sequence 155, Appl
33	70.8	2.9	1056	4	US-09-364-230-27	Sequence 27, Appl
34	70.2	2.9	1230	4	US-09-252-991A-14780	Sequence 14780, A
35	66.2	2.7	1182	4	US-09-252-991A-6534	Sequence 6534, Ap
C 36	66.2	2.7	1419	4	US-09-252-991A-6499	Sequence 6499, Ap
C 37	60	2.4	513	4	US-09-252-991A-9318	Sequence 9318, Ap
38	57.6	2.3	1149	4	US-09-328-352-724	Sequence 724, App
39	54.8	2.2	276	4	US-09-252-991A-9426	Sequence 9426, Ap
40	53.8	2.2	958	4	US-09-155-183-10	Sequence 10, Appl
41	53.8	2.2	1227	4	US-09-328-352-1449	Sequence 1449, Ap
42	53.4	2.2	766	4	US-09-364-230-33	Sequence 33, Appl
43	53.4	2.2	1200	4	US-09-328-352-234	Sequence 234, Appl
C 44	52.4	2.1	7218	1	US-08-232-463-14	Sequence 14, Appl
45	50.8	2.1	1302	4	US-09-252-991A-14660	Sequence 14660, A

ALIGNMENTS

RESULT 1

US-09-252-991A-11446/c
; Sequence 11446, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 11446

; LENGTH: 1341

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11446

Query Match	6.9%	Score 168.6;	DB 4;	Length 1341;
Best Local Similarity	49.9%	Pred. No. 2.9e-37;		
Matches	553;	Conservative 0;	Mismatches 529;	Indels 27; Gaps 4;
Qy	308	AGGTGGACTCCGAAAAAATTCACCAAGGAGGAAATCCCGACATGAACTTTGGAGAAAT 367		
Db	1220	AGCGCGCGCGCGCAGATCGACGAGCAATCAGTTCCCATGGACATGTGGCGCAAGT 1161		
Qy	368	TGAAGAGCCTAGGCGCTTTTGGCGTGCAGTCCAGAGAAATATGTCGCTGGCGTTCT 427		
Db	1160	TCGGCGAGATGGCGCTGCTCGGCATCACCGTCGACGAGGATGCGCGTTCGGCGCTGG 1101		
Qy	428	CCAAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGG---ATGGGTCCATCAGTG 484		
Db	1100	GCTACCTGGCCACGCGCGTGGTTCATGGAAGATACAGCGCGGCTCGGCTCGGTGGCGC 1041		
Qy	485	TGACCTGGCGCGCGCGCGCGCTATTGGCGCTCAAGGGGATCATCTTGGCTGGCGACTGAGG 544		
Db	1040	TCCTCTATGGCGCGCATTCACACCTCTGGTCAACCATGATCAAGCGCAACGACGCGG 981		
Qy	545	AGCAGAAAGCCAAATACCTTGCCTAACTGGCGTCGGGGAGACATTCGAGCTTCTGCGC 604		
Db	980	AAACAGAGCGCGCTACCTACCTCGCGGCGCTGATCCGGGCAACACATCGGTGGCTGGCGA 921		
Qy	605	TCACGAGCGCGCGCTGCGGAGGATGACGCTCAATCCGAGAGACGACACACATGAGTG 664		
Db	920	TGAGCGAACCCACCGCGCGCTCCGAGTGGTCTCGATGAAGCTGCGCGCGCGCGGTCG 861		

SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 5.8%; Score 143.4; DB 3; Length 4403765;
Best Local Similarity 48.2%; Pred. No. 3.3e-28;
Matches 513; Conservative 0; Mismatches 531; Indels 21; Gaps 3;

326 TTGACGAGGAGGAGAAATCCAGATGAAATTTTGAGAAATTTGAAGACCTAGGCGTTT 385
Db TGGACGAGAGCGCGGTTCCCGAAGAGCGCTAGTGGCGCTCAATTCCTCCGTTTCA 3652755

386 TTGGGCTGCAAGTCCGAGAGAAATATGTGGCTGGGCTTCTCCACACCATGACTCA 445
Db ACGCGCTCCACATTCGGAGGAGTACGGCGGTTCAGGGTCCGACTTCGGTAGCTACCTGCA 3652695

446 GACTAGGAGAGATCATCAGCATGGTGGTCCATCCTGTGACCTGCGAGCGCACCAAG 505
QY TCGTGATCGAAGAGGTGGCCCGCTCGACGCTCGCTGCTGATGCCCGGGTCAACA 3652635

506 CTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGAGAGAGAGAAACCAATACTTGC 565
Db AGCTGGACCATGGGCTGATCTTGGGGCTCCGAGAGCTGAAGAGCAGGTGCTGC 3652575

566 CTAACTGGCTCCGGGAGACATTCGACGCTTCTGCTCAGGAGCGACGACCTGGGA 625
QY GTGACCGGGTCCATCGGACGCGGGCC-----AAGCCGATGGGATCATCGATTTC 3652461

626 GCGATCGACGCTCAATCCGAGCAGAGCCACACTAAGTGAAGACAAAGACACTACATCC 685
Db TCAATGGCTCAGAGTCTGGATTACTAATGGAGGACTGCCAATATTTTACTGTGTTTG 745

686 TCAACGGCGCAAGTCTGGATACCAACGGGGGCAAGTGCAGCTGGTACACGGGTATGG 3652401
QY CAAAGACTGAGTCTGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 805

806 TAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAGATATAATTTAGCATTCGG 865
Db TGCAAGAGCAGACGAGGGTTTCCACGCTCGGTCCGAAAGACCAAGCTCGGGATCAAG 3652293

866 GCTCCACACTTGTGAAGTCCATTTTGAAGAACCAAGATACCTGTGGAAACATCCTTG 925
QY GAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATTCCTCAACAGCGCGCGGTTCAG 985

926 GAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATTCCTCAACAGCGCGCGGTTCAG 985
Db GTGAGCGCGTACCGGTTTCAAGACCGCGCTGGCCACAGTGGACACACCCGTCACCA 3652173

986 TGGCAGCGTCTGGCTGGGCTGCTCAAGAGATTGATGAAATGATGCTGTGAGTACGCTT 1045
Db TTGGCGCGCAGCGCGTGGGTATCGCCAGGCGCGCTGGATGCTGCATCGCTACACCA 3652113

1046 GCACAAGGAAACAGTTTAAACAAGAGGCTCAGTGAATTTGATTTGATTCAGAGAGAA 1105
Db AGGACCCCAAGCAGTTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3652112

1106 CACTGATGGCTCAGAAGCTTACGTATGGAGAGTATGACCTACCTCAGACGAGGATGC 1165
QY CCGACATGGGATGAAGTGGAGGCGGCTCGGCTGATGGTGTACTCGCGCCGCCCC---C 3651996

1166 TGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGCGAGCCATGGTGAAGGTGTTCACT 1225
QY TGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGCGAGCCATGGTGAAGGTGTTCACT 1225

580 GGGAGACACATTCAGACCTTCTCCCTCAGGAGCCAGCCAGTGGAGCGATCGACGCTCA 639
Db GGGCAACTGCTCAGTTCTGTTCTGCTTCCCTCAGGAAACGAGTCCGGTTCCGACCGCCCTCG 769

640 ATCCGGAGCAGACCACTAAGTGAAGACAAAGACACTACATCTCAATGGCTCCAAAG 699
Db CTGAAGACCAACCGGCTGCT-----GTGAGCGGACGACTACGTCTCAACGTTACCAAG 715

700 GTCTGATTACTAATGAGGAGCTGGCAATATTTTACTGTGTTTTCGAAAGACTAGGTC 759
Db GCGATCGCTGGCGCC-----GGCGGAATCTCCGCTTTCTCGTTCGAGCGCGGTAGC 604

820 GGTGAGTCTAATATGGAACCCGAAGATAAATAGGCAATTCGGGGCTCCAACTTGT 879
Db CCCGGCTGCTCCCTGGCAAGCCGATCGCAAGATGGACAGAGGCGCGCATACCTGT 544

880 GAAGTCCATTTGAAACACCAAGATACCTGTGGAACATCTTGGAGAGTCT---GGA 936
Db GACGTGATCTTCGACGACTGCGGGTGCCTGCCAGTCACTGATGATCGGGGCTCGAGGGC 484

937 GATGGTTTAAAGTGGCCATGAACATCTCAACAGCGCGGTTTCAGCATGGGCGAGCTC 996
Db GTGGCTGGCTGCTCAAGAGATTGATTGAAATGACTGCTGAGTACGCTCGACAAAGAA 1056

423 TGTGTTGGGTCGCGAGCGGATGCTTGAGGATGCTTGCGCTATGCGCTGGAACGCAAG 364
QY CAGTTTACAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACTGATGGCT 1116

363 CAGTTTCCGCGAGCGATGCGCGATTCAGTTGATCCAGGGGATGCTTGGCAGCAAG 304
QY CAGAGGCTTACGTCATGAGAGTATGACCTTACCTCAGCAGGAGGATGCTGGACCAACT 1176

303 GCGAAGCCTACGGCGCGCTGCTGATGCTATCAGCGCGCGCGCGAGCGCGAGAA--- 247
QY GCTTTCGCTGCTCAGCAGGAGCGGATGAGGTTGAGTTCAGCTCCGAGCGCGCC 1236

246 ---GGCCGCGATGCTGAGTACGAGCGGCTGCGCGAGCTGCTGCGCTCGGAATGTC 190
Db TGCGAGTGTGAGTGGGCTGCGAGATCCTCGGGGCTTGGGCTACACAAAGGAGTAT 1296

189 GCGCGCTGCGCGAGCGCGGTCAGATATTCGGCGCGCGGCTATATCGCGACTAC 130
QY CGGTACGAGCGGATGCTGCTGAGCAGCGGATCCTCCTCATCTTCGAGGGAACCAATGAG 1356

129 GGGATCGAGCGCTTCTATCGGAGCTGCGGCTGTTCCGCATCTACGAGGGCACCACCGAG 70
QY ATTCTCCGATGATACATCGCC 1378

69 ATCCAGCACTGCTATCGCC 48
Db

RESULT 5
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6284328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2

Db 3651995 GCGCCGAACGCGGTGAGCGGATCTGGGCTTCATTTCCGGCGGCTCGAAGTGTCTCGCCT 3651936
QY 1226 CCGAGCGCCCTGGCGAGTGTGTGAGTGGCGCTGCAGATCCTCGGGGCTTGGGCTACA 1285
Db 3651935 CCGAGCTCGGATGGAGTACCACCGATCGGTACAACTGTTCGGGGGCGCGCTACA 3651876
QY 1286 CAGGGGACTATCCGTACGAGCGCATCTGCGTGACACCGCATCTCTCTCATCTTCGAGG 1345
Db 3651875 CCACGACTTCCGCTCGAGCGGTTTATGCGCGAGCGCAAGATCACCCATCTATGAGG 3651816
QY 1346 GAACCAATGAGATCTCCCGATGTACATCGCCCTGACGGGTCTGC 1390
Db 3651815 GCACCAATCAGATTCAGCGGCTAGTGATGTCGGGCGCTACTGC 3651771
RESULT 6
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
Query Match 5.8%; Score 143.4; DB 3; Length 4411529;
Best Local Similarity 48.2%; Pred. No. 3.3e-28;
Matches 513; Conservative 0; Mismatches 531; Indels 21; Gaps 3;
QY 326 TTGACGAGAGGAAATCCAGATGAACTTTTGAGAAATGGAAGAGCTAGGCTTT 385
Db 3657964 TGGACGAGAAGCGCGGTTCCCGGAAGAGCGCTAGTGGCGCTCAATTCCTCCGGTTTCA 3657905
QY 386 TTGGGCTCAACTCCAGNAGATATGTTGGCTTCTCCCAACACCATGACTCAA 445
Db 3657904 ACGCCGTCACATTCGGAGGAGTACGGCGTCAAGGTGCGGACTCGGTAGCTACCTCA 3657845
QY 446 GACTAGGGGAGATCATCAGCATGGATGGTCCATCATCTGTACCCCTGCACCGCACAGG 505
Db 3657844 TCGTGATCGAAGAGTGGCCCGCTGACGGCTCGGCTGCTGTGATCCCGCGTCAACA 3657785
QY 506 CTATTGGCTCAAGGGATCATCTTGGCTGGCACTGAGAGCAGAGCAAAATCTTGC 565
Db 3657784 AGCTGGGACCATCGGCTGTATCTCTCGGGCTCCGAGGAGCTGAAGACAGTGTGC 3657725
QY 566 CTAAACTCGGCTCGGGGAGCACATTCAGGCTTCTGCCTCAGGAGCGCACAGTGGGA 625
Db 3657724 CGCGGTTGGCGCGGAGGCGGATGGCTTCTATGATTTAGTGAACGAGCGCGGCA 3657665
QY 626 GCGATGACGCTCAATCCGAGCAGAGCCACCTAAGTGAAGACAAAGCACTACATCC 685
Db 3657664 GTGACGCGGCTCCATCGGAGCGCGGCC-----AAGCGGATGGGATCACTGGATTC 3657611
QY 686 TCAATGGCTCAAGGCTGATTAATGAGGACTGGGCAATATTTTACTGTGTTG 745
Db 3657610 TCAAGCGCGGCAAGTGTGATCACCAGCGCGCAAGTGCAGCTGTGTACACCGTATG 3657551
QY 746 CAAAGACTGAGGCTGTTGATCTTGTGATGATCAGTGAAGACAAATCAGCAATTCATAG 805
Db 3657550 CGGTGACCGA-----TCCCGACCGGGCGGCGCAACGGCATCTCGGGCTTCATGG 3657503

QY 806 TAGAAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAAAGATAAATTAGCATTCGGG 865
Db 3657502 TGCACAAAGACGACGAGGGGTTTCCCGTGGTCCGAAAGACGCAAGCTCGGATCAAGG 3657443
QY 866 GTTCCAAACACTTTGTGAAGTCCATTTTGAANAACCAACAGATACCTCTGTGAAAACATCCTTTG 925
Db 3657442 GGTACCCGACCCAGCGGTGTACTTCGAGAACTGCCGATCCCCGGCGATCGCATCATCG 3657383
QY 926 GAGAGTCGAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGTTTCAGCA 985
Db 3657382 GTGACCCGCTTACCGGTTTCAAGACCGCGCTGGCCACTTGGACACACACCGTCCCAGCA 3657323
QY 986 TGGGAGCGTCTGTGGCTGGCTGCTCAAGAGATTGATTGAATGACTCTGAGTACGCT 1045
Db 3657322 TTGGCGCGCAGCGCGTGGTATCGCCAGGCGCGCTGGATGCTGCCATCCCTACACCA 3657263
QY 1046 GCACAAGAAACAGATTTTAAAGAGGCTCAGTGAATTTGGATTGATTTCAGGAGAAATTTG 1105
Db 3657262 AGGACCGCAAGCAGTTCGGTGTGATCGATCAGTACGTTTCAGGCGGTGCTGCTG 3657203
QY 1106 CACTGATGCTCAGAAGGCTTACGTCATGAGAGATGATGACTACCTACCTCACAGCAGGATGC 1165
Db 3657202 CCGACATGGCATGAAGGTGGAGCGGCTCGGCTGTGTTTACTCCGCGCGGCC---C 3657146
QY 1166 TGGACCAACCTGGGCTTTCCCGACTGCTCCATCGAGGAGGATGTTGAAGGTGTTCACT 1225
Db 3657145 CGCGCAAGCGGTGAGCGGATCTGGGCTTCAATTCGCGGCTCGAAGTGTTCGCT 3657086
QY 1226 CCGAGGCGCGCTGCGAGTGTGTGAGTGGCGCTGAGAGTCCCTCGGGGGGTTGGGTACA 1285
Db 3657085 CCGAGCTCGGATGGAGTCAACACCGATCGCGTACAACTGTTTCGGCGCGCGGCTACA 3657026
QY 1286 CAAGGACTATCCGTACGAGCGCATCTGCGTGCACACCCGATCTCTCTCATCTTCGAGG 1345
Db 3657025 CCACGACTTCCCGGTGCGAGCGTTTATGCGCGACGCAAGATCACCCAGATCATGAGG 3656966
QY 1346 GAACCAATGAGATCTCCGATGATACATCCCTGAGCGGTCTGC 1390
Db 3656965 GCACCAATCAGATTACGCGGTAGTGTGCGCGCGCTACTGC 3656921

RESULT 7

US-09-221-017B-824/c
; Sequence 824, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:

Tue Sep 9 10:14:36 2003

APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 824:
SEQUENCE CHARACTERISTICS:
LENGTH: 3515 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE: PORYPHOMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...3515
IS-09-221-017B-824

Query Match	5.7%	Score 139.4;	DB 4;	Length 3515;
Best Local Similarity	48.7%	Pred. No. 7.9e-29;		
Matches 516; Conservative	0;	Mismatches 516;	Indels 27;	Gaps 4;

QY	321	AAAAATTGACCGAAGGAAAAATCCCATGAACTTTGGAGAAATTTGAAGAGCCTAGG	380
DB	1655	AGAAATTTGACGACGAGGAACGTTTCCCATCGAAACAGTAAAGAAGATGGCGAAATCGG	1596
QY	381	GCATTTTGGGCTGCAGTGCACGAAGAATATGTCGTGCCCTGGGCTTCTCCAAACACCATGTA	440
DB	1595	ACTGATGGCATCCCTTTTCCCGCTTGAATACGGTGTGCGCGGGTACAAACGTAATGTA	1536
QY	441	CTCAAGACTAGGGAGATCATACGATGGATGGTCCATCACT--GTGACCCCTGCGACG	497
DB	1535	TTCAATGGCAGTGAGGAACATAGCCGTGTATGTCTACCACAGGCGTGGTCTCTCTGC	1476
QY	498	GCACACGCTATTGGCTCAAGGGATCATCTTTGGTGGCACTGAGGACGAGAAACCCAA	557
DB	1475	ACACACATCTTTGTGCGCAGCTCCCATCTACGAAATGGAACAGAAAGAAATGAA	1416
QY	558	ATACTTGGCTAAACTGGCGTCCGGGAGCACATTGCAGCCTCTGCCTCACGAGGCACG	617
DB	1415	GTATTTGCGGAAACTCTGTCTGCGGAATGGATCGGTGCATTTGGACTTACCGAACCCAA	1356
QY	618	CAGTGGGAGCGCATCAGCGCTCAATCCGAGCAGAGCCACACTAAGTGAAGACAGAAGCA	677
DB	1355	TGCAGGTACGATGCTGCTGCTCAGCAAACTTTTGCTGAAGAAAGGAAAG-----ATCA	1302
QY	678	CTACATCTCAATGGCTCCAAAGTCTGGATTACTAATGGAGGACTGCCCAATATTTTAC	737
DB	1301	TTTCTGCTTAACGGTAAACAAGATTTTATCACCAAATGCCGAATATGCTCACCGTATACGT	1242
QY	738	TGTGTTTCCAAAGACTGAGGTCGTTGATTTCTGTGATCGATCAAGACACAAAATCACAGC	797
DB	1241	GGTATTTGCCATGACAGA-----TAAGAGCCAGGCACACGTGGTATCACCGC	1194
QY	798	ATTCTAGTGTAGAAAGAGACTTTGGTGGAGTCACCTAATGGGAAACCCGGAAGATAAATAGG	857
DB	1193	TTTTCATCGTTGAAGAAGAACTCCCGGGTTCTCTATCGAAGAAGAAAGAGCTTAAAGTGGG	1134
QY	858	CATTCGGGCTCCACACTTTGTAAGTCCATTTTGAACACACCAAGATACCTGTGGGAAA	917
DB	1133	TATCCGTGTCGGCTACATCGCACTTATTTTTGAAACACTGTATCTGCCCAAGAGAA	1074

RESULT 8

```

RESULI 8
US-09-252-991A-16537
  Sequence 16537, Application US/09252991A
  Patent NO. 6551795
  GENERAL INFORMATION:
  APPLICANT: Marc J. Rubenfield et al.
  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
  TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
  FILE REFERENCE: 107196.136
  CURRENT APPLICATION NUMBER: US/09/252,991A
  CURRENT FILING DATE: 1999-02-18
  PRIOR APPLICATION NUMBER: US 60/074,788
  PRIOR FILING DATE: 1998-02-18
  PRIOR APPLICATION NUMBER: US 60/094,190
  PRIOR FILING DATE: 1998-07-27
  NUMBER OF SEQ ID NOS: 33142
  SEQ ID NO 16537
  LENGTH: 1395
  TYPE: DNA
  ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16537

```

Query Match	5.5%	Score 134;	DB 4;	Length 1395;
Best Local Similarity	48.8%	Pred. No. 1.5e-27;		
Matches 430:	Conservative	0;	Mismatches 440;	Indels 12; Gaps 2;

[illegible]

Db 686 GTAGCAACAGCTTCATCAGCCACGCGCGGCACCGCGACTTCGCCACTCGTTCGCCGTCA 745
Qy 752 CTGAGGTCTGTGATTTCTGATGATCAGTGAAGACAAATACACAGCATTCATAGTAGAAA 811
Db 746 CCGACAGCTACGACACACGCGCGAAGCGCAACGCGGTGACCGCTTCTCTGTGTGACA 805
Qy 812 GAGACTTTGGTGGAGTCACTAATCGGAACCCGGAAGATAAATTAGGCATTCGGGGCTCCA 871
Db 806 AGGCGACGCGCGGATACCGTGCCTCGCGCGCGGAATCGGTGAGCAACCGTGGCTACC 865
Qy 872 ACATCTGTGAAGTCCATTTTCAACACACCAAGATACCTGTGGAAACATCCTTTGGAGAGG 931
Db 866 ACACCTAGGAGATCTTCTGACGATTTGCCGGGTGCGCGCTCCAGGTCTCTCGCGGAGG 925
Qy 932 TCGGAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGTTTCAGCATGGGCA 991
Db 926 TCGCAAGGGCTGGGAAGTGGCCAAACCGCTGCGTACCAGCGCGGAGGATGGTGGCGG 985
Qy 992 GCGTCTGGTGGCTCTCAAGAGATTGATTGAATGACTGCTGAGTAGCCTGCGACAA 1051
Db 986 CCAACTGTGTGCGGCCAGCGCGCGCGCTCGACCTGCTGCGTGGCGCGCGGATC 1045
Qy 1052 GGAACAGTTTAAACAGAGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACTGA 1111
Db 1046 GCAAGCAGTTTCGCGCAGCGCATCGGAGCTACCGAGCGGCTTCTTCAAGCTCGCGGACA 1105
Qy 1112 TGCTCTAAGAGGCTTACGTATGAGAGATGATGATGAAATGACTGCTGAGTAGCCTGCGACAA 1171
Db 1106 TGCCACAGCAGATCTCGTGGCGCGAGCTGATGACCTGCGACCGCTGGAAGATGGACC 1165
Qy 1172 AACCTGGCTTTCGCGACTGCTCCATCGAGCGCCATCCTCGGGGGCTTGGGCTACACAGGG 1231
Db 1166 AGGCGACCATGACCGACGG-----CGAGGCGCGCATGGCCAAAGCTGTTCCGCGAGGAGA 1219
Qy 1232 CCGCTTGGCAGTGTGTGAGTGGCGCTCGAGATCCTCGGGGGCTTGGGCTACACAGGG 1291
Db 1220 CCGTGGCAAGTTCGCGCAGAGCGGTGTCAGATCTTCGGCGCATGGGCTGTGAGT 1279
Qy 1292 ACTATCCGTACGCGCATCTGCTGACACCGCATCTCTCATCTTCGAGGGAACCA 1351
Db 1280 AAGGACCGTTCGAGCGCATCTGGCGCAAGCGCGGATCGAAGGATCTGGGAGGCACTT 1339
Qy 1352 ATGAGATTCCTCGGATGTACATCGCCCTGACGGGTCTGACG 1393
Db 1340 CGGAATCCAGCGCACATCGTTCCCGGAACTCGTGGCG 1381

RESULT 9

US-09-252-991A-16061/c
; Sequence 16061, Application US/09252991A
; Patent No. 6531795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16061
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16061

Query Match 5.5%; Score 134; DB 4; Length 1665;
Best Local Similarity 48.8%; Pred. No. 1.7e-27;
Matches 430; Conservative 0; Mismatches 440; Indels 12; Gaps 2;

Qy 512 GCTCAAGGGGATCATCTTTGGCTGGCAGTACAGAGCAAGAAACCCAAATACTTTCCTTAAC 571
Db 1364 GGCCTTCGAAGATCTCTATGGCTGCACCGGAGAACAGCTCGGCGACTACTCTTGTGGCCT 1305
Qy 572 TGGCGTCCGGGAGACATTTGAGGCTTCTGCTTACGAGGAGCCAGCCAGTGGAGCGATG 631
Db 1304 GCGTGCAGGCGGAAAAGACCGATTTCTGCGCTTACCGAGCGCGCGCTCCGAGG 1245
Qy 632 CAGCCTCAATCCGGAGCAGCAGCACAATAGTGAAGACAAAGACACTACATCTCTCAATG 691
Db 1244 CCAATTCGATCAAGACCCCGCGGTGCGGAGCGGCGC-----CTTCGTGATCAACG 1191
Qy 692 GCTCAAGGCTGTGAGTACTAATGAGGAGCTGGCCAATATTTTACTGTGTTTGCAGAA 751
Db 1190 GTAGCAAGCACTTTCATCAGCCAGCGCGGAGCTTCGCCATCTCTTCGCCGTCA 1131
Qy 752 CTGAGGTCGTGTGATTTCTGATGATCAGTGAAGACAAATACAGCAATTCATAGTAGAAA 811
Db 1130 CCGCAGCTACGAGCACACCGCGCAAGCGGTCACCGCTTCTGTTGGTGGACA 1071
Qy 812 GAGACTTTGGTGGAGTCACTAATGGAAACCCGAGATAAATAGGCATTCGGGCGCTCCA 871
Db 1070 AGGCGACCCCGGATGACCTGCGCGCGCGCGCGAATGCGTGAGCAACGCTGGCTACC 1011
Qy 872 ACATCTGTGAAGTCCATTTTCAAAACACCAAGATACCTGTGGAAGAAATCTTGGAGAG 931
Db 1010 ACACCTAGAGATCTTCTTCAGCATTTCCCGGCTGCGGCTTCCAAAGTGTCTCGCGAGG 951
Qy 932 TCGGAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGCTTACGATGGGCA 991
Db 950 TCGCAAGGGCTGGGAAGTGGCCAAACCGCTGCGTACCGCGGAGCGGTGATGTTGCGCG 891
Qy 992 GCGTGTGGCTGGCTGCTCAAGAGATTGATTGAATGACTGCTGAGTAGCCTGACAA 1051
Db 890 CCAACTGTGTGCGCGCAGCGCGCGCTCGACCTGCTCGCTGCGCTGGCGCGCGATC 831
Qy 1052 GGAACAGTTTAAACAGAGCTCAGTGAATTTGGATTGATTGAGGAGAAATTTGCACTGA 1111
Db 830 GCAAGCAGTTTCGCGCACCCGATCGGAGCTACGAGGCGGTTTCTTCAAGCTCGCGGACA 771
Qy 1112 TGGCTCAGAAGGCTTACGTCTATGAGAGTATGACTACCTACCTACAGCAGGAGATCTGGACC 1171
Db 770 TGGCAGCAGATCCGTGCGCGCGAGCTGATGACCTTGCACACCGCTGGAAGATGGACC 711
Qy 1172 AACCTGGCTTTCGCGACTGCTCCATCGAGCAGCCATGGTGAAGTGTTCAGTCCGAGG 1231
Db 710 AGGCGACCATGACCGACGG-----CGAGCGCGCATGGCCAAAGCTGTTCCGCGAGGAGA 657
Qy 1232 CCGCTTGGCAGTGTGAGTGAAGGCGCTGAGAGCTCTCGGGGCTTGGGCTACACAGGG 1291
Db 656 CCGTGGCAAGGTCGCGCAGCAGCGGTGCAGATCTTCGCGGCGATGGGCTGTGAGTGG 597
Qy 1292 ACTATCCGTACGAGCGCATCTGCTGACACCGCATCTCTCTCATCTTTCGAGGGAACCA 1351
Db 596 AAGACCGGTTCGAGCGCATCTGCGCGCAACGCGCGGATCGAAGGATCTCGGAGGCGACTT 537
Qy 1352 ATGAGATTCCTCGCGATGTACATCGCCCTGACGGGTCTGCGAGC 1393
Db 536 CGGAATCCAGCGGCACATCGTTTCCCGCGAACTGCTGCGCG 495

RESULT 10

US-08-311-731A-121
; Sequence 121, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:

QY 1165 CTGACACACCTGGCTTTCCGACCTCTCCATCGAGGAGCCATGGTGAAGTGTTCAGC 1224
 Db 225 CGCGAA-----GCCGGCTTGGCTGCTGACGAGCCTCGATGCCAAGCTGTTCGCC 172
 QY 1225 TCCGAGCCGCTGGCAGTGTGTGAGTGGAGGCTGCAGATCCTCGGGGGCTTGGGCTAC 1284
 Db 171 TCGAGATGCCCAGGAGTTTCTCCGCGGCGATCCAGACCTCGGTGGCTATGGCTAC 112
 QY 1285 ACAAGGACATCCGTACGACGATCTGCGTGGTACACCCGATCCTCTCATCTTCGAG 1344
 Db 111 CTCAAGGACTTCCCGGTGGAACATCTATCGGAGCGTGGGGTCTGCCAGATCTACGAG 52
 QY 1345 GGAACCAATGAGATTTCTCGGATGATACATCGCCC 1378
 Db 51 GGCACGAGCGATGCGAGCCCTGGTGTATCGCC 18

RESULT 13

US-09-648-004-5
 ; Sequence 5, Application US/09648004
 ; Patent No. 6498242
 ; GENERAL INFORMATION:
 ; APPLICANT: CHEN, QIONG
 ; APPLICANT: THOMAS, STUART
 ; APPLICANT: NAGARAJAN, VASANTHA
 ; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
 ; TITLE OF INVENTION: INTERMEDIATES
 ; FILE REFERENCE: CL-1341-A
 ; CURRENT APPLICATION NUMBER: US/09/648,004
 ; CURRENT FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/252,553
 ; PRIOR FILING DATE: 1999-02-19
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 5
 ; LENGTH: 1155
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter sp.
 US-09-648-004-5

Query Match 4.5%; Score 111.2; DB 4; Length 1155;
 Best Local Similarity 47.8%; Pred. No. 3.5e-21;
 Matches 499; Conservative 0; Mismatches 518; Indels 27; Gaps 5;

QY 341 AAATCCAGATGAATTTGGAGAAATTTGAAGAGCTAGGCTTTTGGGCTGCAAGTCC 400
 Db 104 AAATCCAGCTGAATTCGTGACGAAATGAAGAACTGGGTCTTTTGGTCTCACCATT 163
 QY 401 CAGAAGATATGGTGGCTGGCTTCTCCA---ACACCATGATCTCAAGACTAGGGAGA 457
 Db 164 CTGAGGAATATGAGGCTTGGCTGACCATGACCATGAGGAGAGGTTTACATTGCAATTGAC 223
 QY 458 TCATCAGCATGGATGGTCCATCACTGTGACCCCTGGCAGCCAGCAGGCTATTGGCCCTCA 517
 Db 224 TGGGAGTACCTCTCTGCTTTCCGTCTACTGCGCACTAACAAATGGATCGGTTCAT 283
 QY 518 AGGGATCATCTTGGCTGGCTGAGGAGCAGAAACCAATATCTTGGCTTAACTGGGCT 577
 Db 284 CAGGCTTAATATTATGATGGCTCCGAGAGCAGAAACAGTATTTTGGCCAGCTCTGGCAA 343
 QY 578 CCGGGAGCATATTGAGCTTCTGCTCAGGAGCCAGCCAGTGGAGCGATGAGCGCT 637
 Db 344 GTGGTGAATATTGTTTCTTCTGTTTAACTGAACCTGATTCCTGGTTCAGATGCTGCT 403
 QY 638 CAATCCGGAGCAGACCACTAAGTGAAGACAAAGCACTACATCTCAATGCTCA 697
 Db 404 CTTT-----AAAAACACAGCGGTGAAGAGTGGTATCATTTACATTTAAATGGCACTA 457
 QY 698 AGGTCTGGATTACTATGAGGACTGGCCAATATTTTACTGTGTTTGAAGAACTGAGG 757
 Db 458 AGCGTTACATCAACCAATGACCCGATGCGGGTGTCTTACTGTCTATGCGCACGTAC----- 512

QY 758 TCGTTGATCTGATGGATCAGTGAAGACAAATTCAGCATTCATAGTAGAAGAGACT 817
 Db 513 ---CAGTACCGAAATTTAAAGCTACAGGTGGAATTTTCAGCCTTTATCGTGACAGTAAAA 568
 QY 818 TTGCTGGAGTCACTAATTTGGGAAACCCGAAAGATAAATTTAGGCATTCGGGGCTTCCAACTT 877
 Db 569 CTCCTGGTATTTCTTGGTTAAAGCTGATGAAGAGATGGGCCAAAGGTGCACATACCT 628
 QY 878 GTGAAGTCCATTTTGAACACCAAGATACCTG---TGGAAACATCTTTGGAGAGTGC 934
 Db 629 GTGATGTATTTTGAACACTGTCTATTTCTGCTATCTGCATCTGCACATTTGGTGGTGTGAAG 688
 QY 935 GAGATGGTTTTAAGGTGCCATGAACATCTCAACAGCGCGGCTTCAGCATGGGACGC 994
 Db 689 GTGTAGGTTTTAAAACTGCAATGAAGTACTTGTATAAGCCGTATTCATATTCCTGCAT 748
 QY 995 TCGTGGCTGGCTGCTCAAGAGATTGATTGAAATGACTGCTGAGTACGCCCTGCACAAGGA 1054
 Db 749 TAAGTGTAGTGTCTGCTACGCTATGCTGGAAGATTCCCTACAATATGCCCTTGAGCGCA 808
 QY 1055 AACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTTCAGGAGAAATTTTGCATGTGG 1114
 Db 809 AACAGTTTGGTCAAGCGATTGCGAACTTCCAGTTGATTCAAGGTATGTTAGCCGATTTCTA 868
 QY 1115 CTCAGAAGGCTTACGCTCATGGAGAGTATGACCTCCTCACAGAGGATGCTGGAACAC 1174
 Db 869 AAGCTGAAATTTACGACGAAATGATGTTAGATTAGCTTGCCTGCTGCTGCTGCTGCTGCTG 923
 QY 1175 CTGGCTTTCCGACTGCTCCATCGAGGAGCCATGCTGAAGTGTTCAGCTCCGAGGCGC 1234
 Db 924 -TGCTGGACAGAAATGTCAGCAGGAGCATCTTGTGCCAAGATGTTTGCCTACTGAAATGT 982
 QY 1235 CTTGGCAGTGTGAGTGGAGGCGCTGCAGATCCTCGGGGCTTGGGCTTACACAAGGACT 1294
 Db 983 GTGGCGGTGTCGAGATCTGCGGTACAGATCCATGTTGTTGCGGTTATATCATGTAAT 1042
 QY 1295 ATCCGTACGAGCGCATACTGCTGACACCCGCGCATCTCTCTCATCTTCGAGGGAACCAATG 1354
 Db 1043 ATGCTATTGAGCGTTTTTACCGTGTATGATGCTTTATTCGTTTGTATGAAGGTACAACGC 1102
 QY 1355 AGATTCCTCGGATGATACATCGCCC 1378
 Db 1103 AAATCCACAGGTCATATTGCCCC 1126

RESULT 14

US-09-648-004-27
 ; Sequence 27, Application US/09648004
 ; Patent No. 6498242
 ; GENERAL INFORMATION:
 ; APPLICANT: CHEN, QIONG
 ; APPLICANT: THOMAS, STUART
 ; APPLICANT: NAGARAJAN, VASANTHA
 ; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
 ; TITLE OF INVENTION: INTERMEDIATES
 ; FILE REFERENCE: CL-1341-A
 ; CURRENT APPLICATION NUMBER: US/09/648,004
 ; CURRENT FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/252,553
 ; PRIOR FILING DATE: 1999-02-19
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 27
 ; LENGTH: 17417
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter sp.
 US-09-648-004-27

Query Match 4.5%; Score 111.2; DB 4; Length 17417;
 Best Local Similarity 47.8%; Pred. No. 1.6e-20;
 Matches 499; Conservative 0; Mismatches 518; Indels 27; Gaps 5;
 QY 341 AAATCCAGATGAATTTGGAGAAATTTGAAGAGCTAGGCTTTTGGGCTGCAAGTCC 400

1880 AAATCCAGCTGAAATCGTCAGCAAAATGAAGAAGCTGGGTCTTTTGGTCTCACCAATC 1939
401 CAGAGAAATATGGTGGCTGGCTTCCCA---ACACATGTACTCAAGACTAGGGGAGA 457
1940 CTGAGGAATATGAGGGCTTGGCTGACCATGGAGAGAGGTTTACATTCGATTGCAAC 1999
458 TCATCAGCATGGATGGGTCCATCACTGTGACCTGGCAGCGCAGGCTATTGGGCTCA 517
2000 TGGGAGGTACCTCTCCTTCCTTCGTTCACTGATCGCACTAACATGGGATCGGTTTCAT 2059
518 AGGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCAAAATACTTGCCTTAAACTGGCGT 577
2060 CAGGCTTAATATTGATGGCTCCGAGAGCAGAAACAGTAGTATTTTGGCCAGCTGGCAA 2119
578 CCGGGGAGCAGCATGACGCTTCTGCTCAGGAGCAGCAGCTGGGAGGATGACGCT 637
2120 GTGGTGAATATTGGTTCAATTCGTTTAACTGAACCTGATCCGGTTTCAGATGCTGCCT 2179
638 CAATCCGGAGAGCCACACTAAGTGAAGACAGAAAGCAAGCACTACATCCTCAATGGCTCA 697
2180 CTTT-----AAAAACCACAGCGGTGAAGATGCTGATCATTAATTTAAATGGCACTA 2233
698 AGGTCTGGATTACTATGGAGGACTGCCCAATATTTTACTGTGTTTGAAGAGCTGAGG 757
2234 AGGTTACATCAACAATGACCGCATGCGGGTCTTTTACTGTGATGCGACGTAC----- 2288
758 TCGTTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2288
2289 ----CAGTACCGAAATTAAGGTACAGGTGGAATTTACAGCCTTTATCGTGGAGCAAAA 2344
818 TTGGTGGATCACTAATGGAAACCCGAAGATAAATTAGGCATTCGGGGCTCCACACTT 877
2345 CTCCTGATTTCTTGGTGAACGTGATAAGAGATGGCCAAAAGGTGACATACCT 2404
878 GTGAAGTCCATTTTGAAACACCAAGATACCTG---TGAAACATCCTTGGAGAGTGG 934
2405 GTGATGTGATTTTGAAGTGTGATTTCTTGCATCTGCATCTCATCTGATGATGATGATG 2464
935 GAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGCGGTTACAGATGGCAGCG 994
2465 GTGATGGTTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 994
995 TCGTGGCTGGCTGCTCAAGAGATTCATTGAAATGACTGCTGAGTACGCTGCAACAGGA 1054
2525 TAAGTGTAGTGTGCTACGCGTATGCTGGAAGATTCCTTACAATATGCCGTTGAGCGCA 2584
1055 AACAGTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACCTGATGG 1114
2585 AACAGTTTGGTCAAGCGATTGGGAATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 2644
1115 CTCAGAGGCTTACGTCATGGAGATGATGACCTACCTCAGCAGGAGGATGCTGGACCAAC 1174
2645 AGCTGAATTTACCGCAGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATG 2699
1175 CTGGCTTTCCGACTGCTCCATCGAGGAGCAGCATGTTGAAGTGTTCAGCTCCGAGGCG 1234
2700 -TGCTGGAGAGATGTCAGCAGCAGGAGCATCTTGTGCCAAGATGTTGGCCACTGAAATGT 2758
1235 CCGTGCATGTGTGATGAGGCGGTGACAGATCCTCGGGGGCTTGGGCTACACAGGAGCT 1294
2759 GTGGCGGTGTCGCGATCGTGGCGTACAGATCCATGTTGTTGGCGGTTTATATCATGTAAT 2818
1295 ATCCGTACGAGCGATGCTGCTGACACCCCGCATCTCTCATCTTCCAGGGAACCAATG 1354
2819 ATGCTATTGACGGTTTTTACCGTGTATGATGATGATGATGATGATGATGATGATGATGATG 2878
1355 AGATTCCTCGGATGATACATCGCC 1378
2879 AAATCCAACAGGTCAATTATTGCC 2902

RESULT 15

US-09-328-352-2254
; Sequence 2254, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2254
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2254

Query Match 4.3%; Score 105.8; DB 4; Length 1170;
Best Local Similarity 47.3%; Pred. No. 1.2e-19;
Matches 502; Conservative 0; Mismatches 532; Indels 27; Gaps 5;
QY 323 AAATTCACGAGGAGGAAATCCAGATGAACCTTTGGACAAATTCAGAGCCCTAGGCC 382
Db 101 AAGTGTAGAGAGAGACAAATTCAGATGACATTTGCACAAATTCGAGTGGTT 160
QY 383 TTTTGGCTGCAAGTCCCAAGAAATATGTCCTGGCTTCTCCAAACACCATGTA 442
Db 161 TATTTGGCTCACAATTTCCCAAGAAATATGTCCTGGCTTCTCCAAACACCATGTA 442
QY 443 CAAGACT---AGGGAGATCATCAGCATGGATGGTCCATCCTGTCACCTGTCAGCCG 499
Db 221 TCAGATGTGATTTGAATTCAGGACAACTTACCGGCTTTTTCGTTCTTAAATGGTACCA 280
QY 500 ACCAGGTATTTGGCTCAAGGGGATCATTTGGCTGGCACTGAGGAGCAGAAAGCCAAAT 559
Db 281 ATAACGGTATTTGGCTCAAGTGAATTTTGGTATGATGACAGAGAGCAGAAAGCCAAAT 340
QY 560 ACTTGGCTTAACCTGGCTCGGGAGGACATATGTCAGCTTCTGCTCAGGAGCCAGCCA 619
Db 341 ATTTACCACTGTTACGCAAGTGGGCAATTTATGTTTCATTTGTTTAAACGAGCTGAAT 400
QY 620 GTGGAGGATGTCAGCTCAATCCGAGAGCAGAGCAGCTAAGTGAAGACAGCAAGCACT 679
Db 401 CTGTTCTGATGCTGCTCTTTAAACACCACTGCCGTA-----AAAGATGGCGATTTT 454
QY 680 ACATCCTCAATGGCTCAAGGTCTGGATTAATAAGAGCTGGCAATATTTTACTG 739
Db 455 ATGTTATTAACCGCAGCAACGTTTATTAACCAATGACCCCATGCTGCAACTTTTACCG 514
QY 740 TGTTCGAAAGCTGAGTCTGATTCATGATGATGATGATGATGATGATGATGATGATGAT 799
Db 515 TAATGCTCGAC-----CAATCCTGAAATTAAGGGTGGGTGGATTTTACGTT 565
QY 800 TCATAGTAGAAGAGACTTTTGGTGGTCACTAATGGGAAACCCGAAAGATAAATAGGCA 859
Db 566 TTTTATGTCAGGCAATACACCGGCTTACATTTAGGCAAAATTTGACCAAAATTTGGGC 825
QY 860 TCCGGGCTCCCACTTGTGAAGTCCATTTTGAACCAACCAAGATACCTG---TGGAAA 916
Db 626 AAAGGGTTCACATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 685
QY 917 ACATCCTGGAGAGTTCGGAGATGGTTTAAGTGGCCATGAACATCCTCAACAGCGCC 976
Db 686 TGATTTGGTGGCTGAGAGGTGTCGGTTTAAACGCGCATGAAGTCTGGGATAAAGGC 745
QY 977 GGTTCAGATGGGAGCGTCTGGCTGGCTGCTCAAGAGATTTGATTAAGATGACTGCTG 1036
Db 746 GTTTCACATTTGTCCTATAGCTGGGTGTTTGGCAAGCGCATGTTTAAACGAGCATTTGA 805
QY 1037 AGTACGCTGCACAGGAAACAGTTTAAACAGAGCTCAGTGAATTTTGGATTTGATTCAGG 1096
Db 806 ACTATGCGATTGAGCGTAAGCAGTTTGGTCAAGCTTATTTGCGAAGCTTTTTCAGCTTATTCAG 865

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 23:17:14 : Search time 554.612 Seconds
(without alignments)
10168.445 Million cell updates/sec

Title: US-09-945-326-1
Perfect score: 2452
Sequence: 1 cgtgtgtgtcctcgcgc.....actgttaaaaaaaaaaaaaa 2452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1537136 seqs, 1149988732 residues
Total number of hits satisfying chosen parameters: 3074272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2452	100.0	2452	10	US-09-945-326-1
2	2438.4	99.4	2440	14	US-10-168-274-51
3	1863	76.0	1863	10	US-09-945-326-3
4	1102.4	45.0	2080	14	US-10-198-846-11756
5	457.6	18.7	502	11	US-09-918-995-20366
6	452.2	18.4	490	11	US-09-918-995-21901
7	451.8	18.4	483	11	US-09-918-995-23707
8	390.6	15.9	451	11	US-09-917-800A-11034
9	388.6	15.8	2117	10	US-09-918-995-1584
10	261.4	10.7	490	11	US-09-918-995-16927
11	258.6	10.5	312	9	US-09-822-849A-480
12	187.4	7.6	402	11	US-09-918-995-4541
13	165.2	6.7	568	10	US-09-796-692-8175
14	165.2	6.7	568	14	US-10-040-862-8175
15	159.4	6.5	183	10	US-09-867-701-3392
16	151.2	6.2	1158	14	US-10-156-761-5259

c	17	151.2	6.2	9025608	14	US-10-156-761-1	Sequence 1, Appli
	18	147.6	6.0	1170	14	US-10-156-761-6897	Sequence 6897, Ap
	19	147.6	6.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
	20	146.6	6.0	1829	10	US-09-880-107-2309	Sequence 2309, Ap
	21	132.6	5.4	1155	14	US-10-156-761-5010	Sequence 5010, Ap
	22	126.6	5.2	1716	10	US-09-974-300-1101	Sequence 1101, Ap
	23	120.8	4.9	1138	10	US-09-974-300-1082	Sequence 1082, Ap
	24	119.8	4.9	1149	14	US-10-156-761-6577	Sequence 6577, Ap
	25	118.2	4.8	1866	10	US-09-917-800A-1617	Sequence 1617, Ap
	26	115.6	4.7	810	10	US-09-974-300-5530	Sequence 5530, Ap
	27	114.4	4.7	1155	14	US-10-156-761-5264	Sequence 5264, Ap
	28	111.2	4.5	1155	14	US-10-272-419-5	Sequence 5, Appli
	29	111.2	4.5	17417	14	US-10-272-419-27	Sequence 27, Appli
	30	110.6	4.5	1056	14	US-09-974-300-5546	Sequence 5546, Ap
	31	105	4.3	1938	14	US-10-156-761-2554	Sequence 2554, Ap
c	32	104.2	4.2	411	10	US-09-960-352-1781	Sequence 1781, Ap
	33	103.6	4.2	439	10	US-09-998-598-317	Sequence 317, App
	34	102.4	4.2	495	11	US-09-918-995-350	Sequence 350, App
	35	101.8	4.2	439	10	US-09-998-598-317	Sequence 1771, Ap
	36	100.2	4.1	562	10	US-10-060-036-766	Sequence 766, App
	37	94.6	3.9	2238	14	US-09-998-598-1171	Sequence 1171, Ap
	38	92.6	3.8	983	10	US-10-205-823-3	Sequence 3, Appli
	39	90	3.7	2682	10	US-09-974-300-1090	Sequence 1090, Ap
	40	90	3.7	2682	12	US-10-354-358-27	Sequence 27, Appli
	41	86.8	3.5	466	10	US-09-764-847-110	Sequence 110, App
	42	86.8	3.5	466	14	US-10-092-154-110	Sequence 110, App
c	43	84.6	3.5	650	13	US-10-027-632-229219	Sequence 229219,
	44	81.8	3.3	615	10	US-09-974-300-1102	Sequence 1102, Ap
	45	81.4	3.3	1383	10	US-09-925-300-705	Sequence 705, App

ALIGNMENTS

RESULT 1
US-09-945-326-1
; Sequence 1, Application US/09945326
; Patent No. US20020127680A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Hunter, John Joseph
; TITLE OF INVENTION: 62112, A NOVEL HUMAN DEHYDROGENASE AND
; FILE REFERENCE: MN1-187
; FILE REFERENCE: MN1-187
; CURRENT APPLICATION NUMBER: US/09/945,326
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/229,831
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(1932)
US-09-945-326-1

Query Match	100.0%;	Score 2452;	DB 10;	Length 2452;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2452;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	CGTGTGTGTGTCCCTCGCGCGCTAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCG	60	
DB	1	CGTGTGTGTGTCCCTCGCGCGCTAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCG	60	
QY	61	GCACATGAGCGGCTGCGGCTCTTCTGCGCACCGGCTGCGGCTGCGCTGCGCGG	120	
DB	61	GCACATGAGCGGCTGCGGCTCTTCTGCGCACCGGCTGCGGCTGCGCTGCGCGG	120	
QY	121	GCTGTGTGTGTCTTCTACCGGCAACCGGGGCTACTGCGCACCGGCTGCTACGAGCT	180	

Db 121 GGTCTGTGGTCTCTACCGCGAACCGCGGGCTACTCGGCACCGCCGCGCTGTACGAGCT 180
Qy 181 TTGCCCCAAGAGCTTTTCTTAGGCAAAATCAAGAGAAAGAAAGTTTCCCATTTCCAGAA 240
Db 181 TTGCCCCAAGAGCTTTTCTTAGGCAAAATCAAGAGAAAGAAAGTTTCCCATTTCCAGAA 240
Qy 241 GTTAGCCCAAGATGAATTAATGAATCAATCAGTTCTTGGGACCGCTGGAAATTTCTTC 300
Db 241 GTTAGCCCAAGATGAATTAATGAATCAATCAGTTCTTGGGACCGCTGGAAATTTCTTC 300
Qy 301 ACTGAAGAGTGGACTCCCGAAAAATTTGACAGAGAGGAAATCCAGATGAATTTG 360
Db 301 ACTGAAGAGTGGACTCCCGAAAAATTTGACAGAGAGGAAATCCAGATGAATTTG 360
Qy 361 GAGAAATTTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCCTG 420
Db 361 GAGAAATTTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCCTG 420
Qy 421 GGCCTTCCACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATC 480
Db 421 GGCCTTCCACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATC 480
Qy 481 ACTGTGACCCCTGGCAGCGCACAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCACT 540
Db 481 ACTGTGACCCCTGGCAGCGCACAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCACT 540
Qy 541 GAGGAGCAGAAAGCAATACTTCCCTAACTGSCGTCCGGGAGACACATTCAGAGCCTTC 600
Db 541 GAGGAGCAGAAAGCAATACTTCCCTAACTGSCGTCCGGGAGACACATTCAGAGCCTTC 600
Qy 601 TGCCCTCAGGAGCCAGCCAGTGGAGCGATGACGCCCTCAATTCGGAGCAGAGCCACACTA 660
Db 601 TGCCCTCAGGAGCCAGCCAGTGGAGCGATGACGCCCTCAATTCGGAGCAGAGCCACACTA 660
Qy 661 ACTGAAGACAGAAGCACTACATCTCAATGCTCCAGGCTCTGATTAATGGAGGA 720
Db 661 ACTGAAGACAGAAGCACTACATCTCAATGCTCCAGGCTCTGATTAATGGAGGA 720
Qy 721 CTGGCCAAATATTTTACTGTGTTTGCAAAGACTGAGTCTGATGATGATGATG 780
Db 721 CTGGCCAAATATTTTACTGTGTTTGCAAAGACTGAGTCTGATGATGATGATG 780
Qy 781 AAAGACAAATCACAGCATTCATAGTAGAAGAGACTTTGTGGAGTCACTAATGGGAAA 840
Db 781 AAAGACAAATCACAGCATTCATAGTAGAAGAGACTTTGTGGAGTCACTAATGGGAAA 840
Qy 841 CCGCAAGATAAATTAGGCATTCGGGCTCCAAACACTTGTGAAGTCCATTTTGAACAAC 900
Db 841 CCGCAAGATAAATTAGGCATTCGGGCTCCAAACACTTGTGAAGTCCATTTTGAACAAC 900
Qy 901 AAGATACCTGTGAAACATCCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAAC 960
Db 901 AAGATACCTGTGAAACATCCTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAAC 960
Qy 961 ATCCTCAACAGCGCCGGTTACAGATGGCAGCGCTGCTGAGGCTGCTCAAGAGATTG 1020
Db 961 ATCCTCAACAGCGCCGGTTACAGATGGCAGCGCTGCTGAGGCTGCTCAAGAGATTG 1020
Qy 1021 ATTGAATGACTGCTGAGTACCGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAA 1080
Db 1021 ATTGAATGACTGCTGAGTACCGCTGCACAAAGAAACAGTTTAAAGAGGCTCAGTGAA 1080
Qy 1081 TTTGGATTGATTGAGGAAATTTGCACTGATGGCTCAGAAGGCTTACGTCATGGAGAT 1140
Db 1081 TTTGGATTGATTGAGGAAATTTGCACTGATGGCTCAGAAGGCTTACGTCATGGAGAT 1140
Qy 1141 ATGACCTACCTCAGAGAGGATGCTGGACCAACACTGGCTTTCCGACTGCTCCATCGAG 1200
Db 1141 ATGACCTACCTCAGAGAGGATGCTGGACCAACACTGGCTTTCCGACTGCTCCATCGAG 1200
Qy 1201 GCAGCCATGTAAGGTGTTACGCTCCGAGGCGCGCTGGCAGTGTGTGAGTGAAGGCGCTG 1260
Db 1201 GCAGCCATGTAAGGTGTTACGCTCCGAGGCGCGCTGGCAGTGTGTGAGTGAAGGCGCTG 1260

Qy 1261 CAGATCCTCGGGGCTTGGGCTACACAAGGAGACTATCCGTACAGGCGCATACTGCGTGAC 1320
Db 1261 CAGATCCTCGGGGCTTGGGCTACACAAGGAGACTATCCGTACAGGCGCATACTGCGTGAC 1320
Qy 1321 ACCCGCATCTTCATCTTCGAGGGAACCAATGAGATTCTCGGATGTACATCGCCCTG 1380
Db 1321 ACCCGCATCTTCATCTTCGAGGGAACCAATGAGATTCTTCGGATGTACATCGCCCTG 1380
Qy 1381 ACGGTCTGACAGTCCCGCGCATCTGACTACCAAGATCCATGAGCTTAAACAGGCC 1440
Db 1381 ACGGTCTGACAGTCCCGCGCATCTGACTACCAAGATCCATGAGCTTAAACAGGCC 1440
Qy 1441 AAAGTGAACACAGTCATGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCGCAACT 1500
Db 1441 AAAGTGAACACAGTCATGATACCGTTGGCCGAGGCTTCGGGACTCCCTGGGCGCAACT 1500
Qy 1501 GTGGACCTGGGCTGACAGGCAACCATGAGTTGTGACCCCACTCTTCGGGAGAGTGCC 1560
Db 1501 GTGGACCTGGGCTGACAGGCAACCATGAGTTGTGACCCCACTCTTCGGGAGAGTGCC 1560
Qy 1561 AAACAAGTTTGAAGGAACACCTTACTGTTCGGCCGAGCCGCTGGAGACACTGCTCCGC 1620
Db 1561 AAACAAGTTTGAAGGAACACCTTACTGTTCGGCCGAGCCGCTGGAGACACTGCTCCGC 1620
Qy 1621 TTTGSCAAGACCATCATGAGGAGCAGCTGTGACTGAAGCGGTGGCCCAACATCCTCATC 1680
Db 1621 TTTGSCAAGACCATCATGAGGAGCAGCTGTGACTGAAGCGGTGGCCCAACATCCTCATC 1680
Qy 1681 AACCTGTATGGCATGACGCGCGCTGTGTCGGCGGCGAGCCGCTCCATCCGATTGGGCTC 1740
Db 1681 AACCTGTATGGCATGACGCGCGCTGTGTCGGCGGCGAGCCGCTCCATCCGATTGGGCTC 1740
Qy 1741 CGCAACACGACCCAGAGGTTCTTGTGGCCAAACACTTCTGCGTGGAAGCTTACTTTGCGAG 1800
Db 1741 CGCAACACGACCCAGAGGTTCTTGTGGCCAAACACTTCTGCGTGGAAGCTTACTTTGCGAG 1800
Qy 1801 AATCTCTTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTTAGTAGAGAGATT 1860
Db 1801 AATCTCTTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTTAGTAGAGAGATT 1860
Qy 1861 AAGAAAGTGTCCAGCAGATCCTTGAAGAGCAGCTATATCTGTGCCCACTCTGGAC 1920
Db 1861 AAGAAAGTGTCCAGCAGATCCTTGAAGAGCAGCTATATCTGTGCCCACTCTGGAC 1920
Qy 1921 AGGACATGCTGAGGCGAGGACAGTGTCCCTGTCTACCGCCGCGCCCTACCCATGGCCG 1980
Db 1921 AGGACATGCTGAGGCGAGGACAGTGTCCCTGTCTACCGCCGCGCCCTACCCATGGCCG 1980
Qy 1981 TTGCTGGATGACTTGTACTCTTTTTCAGAAAGTGTGGGATTATCAGAGTTAAGCCTT 2040
Db 1981 TTGCTGGATGACTTGTACTCTTTTTCAGAAAGTGTGGGATTATCAGAGTTAAGCCTT 2040
Qy 2041 TTGTTCCCGCTGTGCACCTGAAGGTTGTCCCTGGCCCTGGGAGAGCCCTTCCAGGTTT 2100
Db 2041 TTGTTCCCGCTGTGCACCTGAAGGTTGTCCCTGGCCCTGGGAGAGCCCTTCCAGGTTT 2100
Qy 2101 TGACCTGCAGGAGTGTCTCTACAGGACCATCAGCTTCTGAACCTAGAGCCGAGAGA 2160
Db 2101 TGACCTGCAGGAGTGTCTCTACAGGACCATCAGCTTCTGAACCTAGAGCCGAGAGA 2160
Qy 2161 GAGAAATGGAATTGCTGACCCCTGGAACCTGGGCTGATTTGCTGATTTAGGAGACACCAT 2220
Db 2161 GAGAAATGGAATTGCTGACCCCTGGAACCTGGGCTGATTTGCTGATTTAGGAGACACCAT 2220
Qy 2221 AGTGAACCTGGGCTTATGCTGCTCCAGGCTGAGGTTGGGAGCTGTGTC 2280
Db 2221 AGTGAACCTGGGCTTATGCTGCTCCAGGCTGAGGTTGGGAGCTGTGTC 2280
Qy 2281 AGGTGGAATAGCCATTTCTCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTCT 2340
Db 2281 AGGTGGAATAGCCATTTCTCTCAACACACATTTCTTAAGAAACAGCTTGAAGCTCT 2340

QY 2341 GTCTGGGTCATTTAACTAGAACGAGGAGCTTAAACATGTACAGGAACCAT 2400
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2341 GTCTGGGTCATTTAACTAGAACGAGGAGCTTAAACATGTACAGGAACCAT 2400
QY 2401 TAACAAAGAATATAAATGTGACAACTGTCTACTGTTTAAAAAAGAAAAA 2452
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2401 TAACAAAGAATATAAATGTGACAACTGTCTACTGTTTAAAAAAGAAAAA 2452

RESULT 2
US-10-168-274-51
: Sequence 51, Application US/10168274
: Publication No. US20030124106A1
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: YUE, Henry
: APPLICANT: LAL, Preeti
: APPLICANT: TANG, Y. Tom
: APPLICANT: HILLMAN, Jennifer
: APPLICANT: BAUGHN, Mariah R.
: APPLICANT: AZIMZAI, Valda
: APPLICANT: LU, Dyung Aina M.
: TITLE OF INVENTION: HUMAN OXIDOREDUCTASE PROTEINS
: FILE REFERENCE: PF-0754 PCT
: CURRENT APPLICATION NUMBER: US/10/168, 274
: CURRENT FILING DATE: 2002-08-26
: PRIOR FILING DATE: 1999-12-16
: NUMBER OF SEQ ID NOS: 54
: SOFTWARE: PERL Program
: SEQ ID NO 51
: LENGTH: 2440
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. US20030124106A1 5540437CB1
US-10-168-274-51

Query Match 99.4%; Score 2438.4; DB 14; Length 2440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GTCTCCCTCGGGCGCTAAGAGGGAGACTGAGGCTGAGGCTGGGGACATCGGCAGCA 67
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 GTCTCCCTCGGGCGCTAAGAGGGAGACTGAGGCTGAGGCTGGGGACATCGGCAGCA 60
QY 68 TGAGCGGCTGCGGGGCTCTTCTCGCACACCGCTGCGGCTGCTGCCCTGCGGGGCTG 127
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 TGAGCGGCTGCGGGGCTCTTCTCGCACACCGCTGCGGCTGCTGCCCTGCGGGGCTG 120
QY 128 TGGTCTCTACCGGAACCGCGGCTACTGCGCACCGCGGCTGCGGCTGCGGCTGCGG 187
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 TGGTCTCTACCGGAACCGCGGCTACTGCGCACCGCGGCTGCGGCTGCGGCTGCGG 180
QY 188 AAGAGCTTTCTAGGCAAAATCAAGAGAAAGAGTTTCCCAATTTCCCAAGTAGCC 247
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 AAGAGCTTTCTAGGCAAAATCAAGAGAAAGAGTTTCCCAATTTCCCAAGTAGCC 240
QY 248 AGATGAACCTTAATGAATCAATAGTTCTTGGGACCGGTGGAAAAATTTCTACTGA 307
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 AGATGAACCTTAATGAATCAATAGTTCTTGGGACCGGTGGAAAAATTTCTACTGA 300
QY 308 AGGTGGACTCCCGAAAAATTGACCAAGGAGGAAAAATCCAGATGAACCTTTGAGAA 367
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 AGGTGGACTCCCGAAAAATTGACCAAGGAGGAAAAATCCAGATGAACCTTTGAGAA 360
QY 368 TGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGGCTCT 427
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 TGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGGCTCT 420
QY 428 CCAACACCATCTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCAGTGA 487
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 421 CCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGATGGTCCATCAGCTGTA 480
QY 488 CCCTGGCAGCGGCACACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGC 547
Db 481 CCCTGGCAGCGGCACACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGC 540
QY 548 AGAAGCCAAATACTTGCCTAAACTGGCGTCCGGGAGGACACATTGCAGCCTTCTCGCTCA 607
Db 541 AGAAGCCAAATACTTGCCTAAACTGGCGTCCGGGAGGACACATTGCAGCCTTCTCGCTCA 600
QY 608 CGGAGCCAGCCAGTGGGAGCGGATGAGCCTCAATCCGGAGCAGACCCACACTAAGTGAAG 667
Db 601 CGGAGCCAGCCAGTGGGAGCGATGAGCCTCAATCCGGAGCAGACCCACACTAAGTGAAG 660
QY 668 ACAAGAAGCACTACATCTCTCAATGCTCCAAGGCTCGATTAATCTACTAATCGAGGACTGGCCA 727
Db 661 ACAAGAAGCACTACATCTCTCAATGCTCCAAGGCTCGATTAATCTACTAATCGAGGACTGGCCA 720
QY 728 ATATTTTACTGTGTTTGCAAAGACTGAGGCTGTTGATTTCTGATGGATCAGTGAAGACA 787
Db 721 ATATTTTACTGTGTTTGCAAAGACTGAGGCTGTTGATTTCTGATGGATCAGTGAAGACA 780
QY 788 AATCACAGCATTTCACTAGTAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAG 847
Db 781 AATCACAGCATTTCACTAGTAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAG 840
QY 848 ATAAATAGGCATTCGGGGCTCCAACTTGTGAAGTCCATTTTGAAGAACCAAGATAC 907
Db 841 ATAAATAGGCATTCGGGGCTCCAACTTGTGAAGTCCATTTTGAAGAACCAAGATAC 900
QY 908 CTGTGGAACATCTTGGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCTCTCA 967
Db 901 CTGTGGAACATCTTGGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCTCTCA 960
QY 968 ACAGCGCGGTTTCAGCATGGGAGCGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1027
Db 961 ACAGCGCGGTTTCAGCATGGGAGCGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1020
QY 1028 TGACTGTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087
Db 1021 TGACTGTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1088 TGATTGAGGAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147
Db 1081 TGATTGAGGAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1148 ACCTCACAGGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGGAGCA 1207
Db 1141 ACCTCACAGGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGGAGCA 1200
QY 1208 TGGTGAAGGTTTTCAGCTCCGAGGCGGCTGGCAGTGTGTGAGTGAAGGCGCTGCGAGATCC 1267
Db 1201 TGGTGAAGGTTTTCAGCTCCGAGGCGGCTGGCAGTGTGTGAGTGAAGGCGCTGCGAGATCC 1260
QY 1268 TCGGGGCTTGGCTACACAAAGGAGTATCCCGTACGAGCGCATCTGCTGCTGCTGCTGCTGCTGCT 1327
Db 1261 TCGGGGCTTGGCTACACAAAGGAGTATCCCGTACGAGCGCATCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1328 TCCTCTCTCATCTTTCGAGGAGCAATGAGATCTTCCGGATGTACATCGCCCTGAGCGGTC 1387
Db 1321 TCCTCTCTCATCTTTCGAGGAGCAATGAGATCTTCCGGATGTACATCGCCCTGAGCGGTC 1380
QY 1388 TGCAGCATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1447
Db 1381 TGCAGCATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1448 GCNACTCATGATACCGTTGGCCGAGGCTTTCGGAGTTCCTTGGGCGGCACTGTGGAGC 1507
Db 1441 GCNACTCATGATACCGTTGGCCGAGGCTTTCGGAGTTCCTTGGGCGGCACTGTGGAGC 1500
QY 1508 TGGGGCTGACAGCAGCACTGAGTGTGACACCCAGCTTTCGGGAGTTCCTTGGGCGGCACTGTGGAGC 1567
Db 1501 TGGGGCTGACAGCAGCACTGAGTGTGACACCCAGCTTTCGGGAGTTCCTTGGGCGGCACTGTGGAGC 1560

1568 TTGAGGAGAACACCTACTGCTTTGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGCA 1627
1561 TTGAGGAGAACACCTACTGCTTTGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGCA 1620
1628 AGACCATCATGAGGAGAGCTGGTACTGAAGCGGTGGCCAACTCTCATCAACCTGT 1687
1621 AGACCATCATGAGGAGAGCTGGTACTGAAGCGGTGGCCAACTCTCATCAACCTGT 1680
1688 ATGGCATGACGGCGCTGCTGGCGGCGAGCGCTCCATCCGCAATGGGCTCCCAACC 1747
1681 ATGGCATGACGGCGCTGCTGGCGGCGAGCGCTCCATCCGCAATGGGCTCCCAACC 1740
1748 ACACACAGAGGTCTCTTTGGCGAACACCTTCTGGTGGAGCTTACTTGCAGATCTCT 1807
1741 ACACACAGAGGTCTCTTTGGCGAACACCTTCTGGTGGAGCTTACTTGCAGATCTCT 1800
1808 TCAGCTCTCTCAGCTGGACAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAGAAG 1867
1801 TCAGCTCTCTCAGCTGGACAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAGAAG 1860
1868 TGTCACAGCAGATCTCTTTGAGAGCGAGCTATATCTGTGCCACCTCTGGACAGACAT 1927
1861 TGTCACAGCAGATCTCTTTGAGAGCGAGCTATATCTGTGCCACCTCTGGACAGACAT 1920
1928 GCTGAGCGAGGACAGTGTCCCTGCTACCGCGCGCCCTACCCATGCGCCGCTTGTGG 1987
1921 GCTGAGCGAGGACAGTGTCCCTGCTACCGCGCGCCCTACCCATGCGCCGCTTGTGG 1980
1988 ATGACTCTTACTCTTTTTCAGAGGTGTTGGGATATACAGAGTTAAGCTTTTGTTC 2047
1981 ATGACTCTTACTCTTTTTCAGAGGTGTTGGGATATACAGAGTTAAGCTTTTGTTC 2040
2048 CCGTCTGACCTGAAGGTTGTCGCTGCGCTGGGAGGCTCTCCAGGTTTGTGACCTG 2107
2041 CCGTCTGACCTGAAGGTTGTCGCTGCGCTGGGAGGCTCTCCAGGTTTGTGACCTG 2100
2108 CAGGAGTGTCTCTTAACAGGACCATCACAGCTTCTGAAGTGTGAGCGGAGAGAGATG 2167
2101 CAGGAGTGTCTCTTAACAGGACCATCACAGCTTCTGAAGTGTGAGCGGAGAGAGATG 2160
2168 GAATGCTGACCTGAACTGGCGGTATCTGCTATCTGATGAGGAGACCATAGTGAA 2227
2161 GAATGCTGACCTGAACTGGCGGTATCTGCTATCTGATGAGGAGACCATAGTGAA 2220
2228 ACTGGGCTTATGCTGCTGCTCCAGGCTGTGAGGTGGGTGGGACCTGTGTCAAGTGTG 2287
2221 ACTGGGCTTATGCTGCTGCTCCAGGCTGTGAGGTGGGTGGGACCTGTGTCAAGTGTG 2280
2288 GATAGCCATTTCTGCTCAACACACATCTCTAAGAACAGCTTGAAGCTCTGTCTGG 2347
2281 GATAGCCATTTCTGCTCAACACACATCTCTAAGAACAGCTTGAAGCTCTGTCTGG 2340
2348 TCATTCATTTAACTTAGAGCAGAGGCACTTAAACATGTACACAGGACCATTTAACA 2407
2341 TCATTCATTTAACTTAGAGCAGAGGCACTTAAACATGTACACAGGACCATTTAACA 2400
2408 GAATATAAATGTCACATCTGTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2447
2401 GAATATAAATGTCACATCTGTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2440

Query Match 76.0%; Score 1863; DB 10; Length 1863;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 67 ATGAGCGCTCGGGCTCTTCTGCGCACACGCTGCGGCTGCTGCTGCGGGTCTG 125
Db 1 ATGAGCGCTCGGGCTCTTCTGCGCACACGCTGCGGCTGCTGCTGCGGGTCTG 60
QY 127 GTGCTCTCTACCGGGAACCGCGGCTACTGCGCACACGCTGCGGCTGCTGCTGCGGGTCTG 186
Db 61 GTGCTCTCTACCGGGAACCGCGGCTACTGCGCACACGCTGCGGCTGCTGCTGCGGGTCTG 120
QY 187 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAAAGTTAGC 246
Db 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAAAGTTAGC 180
QY 247 CAAGATGAACCTTAATGAATCAATCAGTTCTTTGGGACCGCTGGAAATTTCTCACTGAA 306
Db 181 CAAGATGAACCTTAATGAATCAATCAGTTCTTTGGGACCGCTGGAAATTTCTCACTGAA 240
QY 307 GAGGTGGACTCCGCAAAATTTGACGAGGAAGAAATCCAGATGAACTTTGGAGAA 366
Db 241 GAGGTGGACTCCGCAAAATTTGACGAGGAAGAAATCCAGATGAACTTTGGAGAA 300
QY 367 TTCAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTTGGCTTTC 426
Db 301 TTCAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTTGGCTTTC 360
QY 427 TCAACACACATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCACTGTG 486
Db 361 TCAACACACATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCACTGTG 420
QY 487 ACCCTGGACGCGACCGAGCTATTGGCTCAAGGGGATCATCTGGCTGGCACTGAGGAG 546
Db 421 ACCCTGGACGCGACCGAGCTATTGGCTCAAGGGGATCATCTGGCTGGCACTGAGGAG 480
QY 547 CAGAAAGCAAAATACCTTGGCTTAACTGGGCTCGGGGAGCAGATTGCGACCTTCTGCTTC 606
Db 481 CAGAAAGCAAAATACCTTGGCTTAACTGGGCTCGGGGAGCAGATTGCGACCTTCTGCTTC 540
QY 607 ACGGAGCGAGCTAGTGGGAGGCTGAGCTCAATCCGAGCAGAGCCACACTAAGTGAA 666
Db 541 ACGGAGCGAGCTAGTGGGAGGCTGAGCTCAATCCGAGCAGAGCCACACTAAGTGAA 600
QY 667 GACAAGAGCAGCTACATCTCAATGGCTCAAGGCTTGGATTTACTAATGGAGACTGGCC 726
Db 601 GACAAGAGCAGCTACATCTCAATGGCTCAAGGCTTGGATTTACTAATGGAGACTGGCC 660
QY 727 AATATTTTACTGTGTTTGCAGAGCTGAGGTGCTGATGATGATGATGATGATGATGATGAT 786
Db 661 AATATTTTACTGTGTTTGCAGAGCTGAGGTGCTGATGATGATGATGATGATGATGATGAT 720
QY 787 AATATCACAGCTTATAGTAGAAGAGCTTTGGTGGAGTCACTAATGGGAAACCCGAA 846
Db 721 AATATCACAGCTTATAGTAGAAGAGCTTTGGTGGAGTCACTAATGGGAAACCCGAA 780
QY 847 GATAAATAGGCAATTCGGGCTCCAACTTTGAGTCACTTTTGAACACCAAGATA 906
Db 781 GATAAATAGGCAATTCGGGCTCCAACTTTGAGTCACTTTTGAACACCAAGATA 840

RESULT 3
US-09-945-326-3
Sequence 3, Application US/09945326
Patent No. US20020127680A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel
APPLICANT: Hunter, John Joseph
TITLE OF INVENTION: 62112, A NOVEL HUMAN DEHYDROGENASE AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: MNI-187
CURRENT APPLICATION NUMBER: US/09/945,326

```
QY 907 CCTGTGGAACATCCTTTGGAGAGCTCGAGATGGGTTTAAGTGGCCATGAACATCCTC 966
Db 841 CCTGTGGAACATCCTTTGGAGAGCTCGAGATGGGTTTAAGTGGCCATGAACATCCTC 900
QY 967 AACAGCGCGGTTGAGATGGGAGCGTCTGGCTGGCTCAAGAGATTCATTGAA 1026
Db 901 AACAGCGCGGTTGAGATGGGAGCGTCTGGCTGGCTCAAGAGATTCATTGAA 960
QY 1027 ATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAAGAGGCTCAGTAAATTTGA 1086
Db 961 ATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAAGAGGCTCAGTAAATTTGA 1020
QY 1087 TTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATATGACC 1146
Db 1021 TTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATATGACC 1080
QY 1147 TACCTCACAGCAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGCC 1206
Db 1081 TACCTCACAGCAGGATGCTGGACCAACCTGGCTTTCCGACTGCTCCATCGAGGAGCC 1140
QY 1207 ATGCTGAAGTGTTCAGCTCCGAGCGCGCTGCGAGTGTGTGAGTGGCGCTGCAAGATC 1266
Db 1141 ATGCTGAAGTGTTCAGCTCCGAGCGCGCTGCGAGTGTGTGAGTGGCGCTGCAAGATC 1200
QY 1267 CTCGGGGCTTGGCTACACAGGACTATCCGTACGAGCGCATCTGCGTGTACACCCCG 1326
Db 1201 CTCGGGGCTTGGCTACACAGGACTATCCGTACGAGCGCATCTGCGTGTACACCCCG 1260
QY 1327 ATCTCTCTCATCTTCAGAGGAACCAATGAGATTCCTGGATGTACATCGCCGTGAGGCT 1386
Db 1261 ATCTCTCTCATCTTCAGAGGAACCAATGAGATTCCTGGATGTACATCGCCGTGAGGCT 1320
QY 1387 CTCAGCATCCCGCGCATCTCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGTG 1446
Db 1321 CTCAGCATCCCGCGCATCTCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGTG 1380
QY 1447 AGCAGAGTCATGATACCTGTGCGGAGGCTTCGGGACTCCCTGGCGGCAACTGTGGAC 1506
Db 1381 AGCAGAGTCATGATACCTGTGCGGAGGCTTCGGGACTCCCTGGCGGCAACTGTGGAC 1440
QY 1507 CTGGGCTGACAGGCAACCATGGAGTTGTGCAACCCAGTCTTGGGAGAGTCCCAACAG 1566
Db 1441 CTGGGCTGACAGGCAACCATGGAGTTGTGCAACCCAGTCTTGGGAGAGTCCCAACAG 1500
QY 1567 TTTGAGGAGAACACCTACTGCTTGGCGGACCGTGGAGACTCTGCTCGCTTTGGC 1626
Db 1501 TTTGAGGAGAACACCTACTGCTTGGCGGACCGTGGAGACTCTGCTCGCTTTGGC 1560
QY 1627 AAGACCATCATGGAGAGCAGCTGTGACTGAAGCGGTGGCCAAACATCTCATCAACCTG 1686
Db 1561 AAGACCATCATGGAGAGCAGCTGTGACTGAAGCGGTGGCCAAACATCTCATCAACCTG 1620
QY 1687 TATGGATGACGGCGTGTGCGGGGCCAGCCGCTCCATCCGATTTGGGCTCCGCAAC 1746
Db 1621 TATGGATGACGGCGTGTGCGGGGCCAGCCGCTCCATCCGATTTGGGCTCCGCAAC 1680
QY 1747 CACGACACAGAGGTTCTCTTGCCCAACACCTTCTCGGTGGAGCTTACTTCGAAATCTC 1806
Db 1681 CACGACACAGAGGTTCTCTTGCCCAACACCTTCTCGGTGGAGCTTACTTCGAAATCTC 1740
QY 1807 TTAGCCCTCTCAGCTGGACAGTATGCTCCAGAAACCTTAGATGAGCAGATTAAAGAA 1866
Db 1741 TTAGCCCTCTCAGCTGGACAGTATGCTCCAGAAACCTTAGATGAGCAGATTAAAGAA 1800
QY 1867 GTGTCCACAGCAGATCTTTGAGAGCGGCTTATATCTGTGCCCCACCCTCTGGACAGACA 1926
Db 1801 GTGTCCACAGCAGATCTTTGAGAGCGGCTTATATCTGTGCCCCACCCTCTGGACAGACA 1860
QY 1927 TGC 1929
Db 1861 TGC 1863
```

RESULT 4

```
US-10-198-846-11756
: Sequence 11756, Application US/10198846
: Publication No. US20030099974A1
: GENERAL INFORMATION:
```

```
: APPLICANT: Lillie, James
: APPLICANT: Xu, Yongyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steinmann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
: TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: TITLE OF INVENTION: THERAPY OF BREAST CANCER
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198,846
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306,220
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11756
: LENGTH: 2080
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 2075, 2076, 2077, 2078, 2079, 2080
: OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11756
```

Query Match

45.08; Score 1102.4; DB 14; Length 2080;

Best Local Similarity 91.9%; Pred. No. 0;

Matches 1359; Conservative 0; Mismatches 81; Indels 38; Gaps 17;

```
QY 26 GAAGGGAGACTGAGGCTGAGGCTGGGGAACATCGGCAGCATGAGCGCTGGGGCTCT 85
Db 450 GAAGGCCAGACTGAGGCTGAGGCTGGGGAACATCGGCAGCATGAGCGCTGGGGCTCT 509
QY 86 TCCTTGGCCACACCGCTGCGGCTGCTGCTGCGGGGTCTGGTGTCTTACCGCAACC 145
Db 510 TCCTTGGCCACACCGCTGCGGCTGCTGCTGCGGGGTCTGGTGTCTTACCGCAACC 569
QY 146 GCGGCTACTTGGCCACCGCTGCTGAGAGCTTTCGCCAAGAGCTTTCTTCTAGGCA 205
Db 570 GCGGCTACTTGGCCACCGCTGCTGAGAGCTTTCGCCAAGAGCTTTCTTCTAGGCA 639
QY 206 AATCAAGAGAAAGAGTTTCCCAATTCAGAAAGTTAGCCAAAGATGAATTAATGAA 265
Db 630 CAATCCCGAAGAG-----GTTAAGTTAGCCCAAGATGAATTAATGAA 674
QY 266 TCAATCAGTTCTTGGACCGCTGGAATAATTTCTTCACTGAGAGTGGACTCCCGAATA 325
Db 675 TCAATCAGTTCTTGGACCGCTGGAATAATTTCTTCACTGAGAGTGGACTCCCGAATA 734
QY 326 TTGACCAGGAA -GGGAAATCCAGATGAACCTTTGGAGAAA -TTCAAGAGCTAGGGCT 383
Db 735 TTGACCAGGAAAGCGGAAATCCAGATGAACCTTTGGAGAAA -TTCAAGAGCTAGGGCT 794
QY 384 TTTTGGGCTCC -AAGTCCCAAGAGATATGTTGGCTGGGCTGGGCTTCT -CCAACACCATGTAC 441
Db 795 TTTTGGGCTCCAAAGTCCCAAGAGATATGTTGGCTGGGCTGGGCTTCTCCCAACCATGTAC 854
QY 442 TCAAGACTAGGGAGATCATCAGATGGATGGTCCATCAGTCCATCAGTCCCTGGCAGCGAC 501
Db 855 TCAAGACTAGGGAGATCATCAGATGGATGGTCCATCAGTCCATCAGTCCCTGGCAGCGAC 914
QY 502 CAGGCTA -TTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGCAAGAACCAATA 560
Db 915 CAGGCTAATTTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGCAAGAACCAATA 974
QY 561 CTTGCTTAACCTGGGCTCGGGGAGACATTCAGACCTTCTGCTTACGGAGCCAGCCAG 620
Db 975 CTTGCTTAACCTGGGCTCGGGGAGACATTCAGACCTTCTGCTTACGGAGCCAGCCAG 1034
```


Best Local Similarity 99.3%; Pred. No. 4e-130;
Matches 454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1415 CCAGATCCATGAGCTTAACAGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCCGGA 1474
Db 33 CCAAGATCCATGAGCTTAACAGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCCGGA 92
QY 1475 GGCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGCGTGACAGGCAACCATGGAGTTG 1534
Db 93 GGCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGCGTGACAGGCAACCATGGAGTTG 152
QY 1535 TGCACCCAGCTTTGGGAGCAGTGCCCAACAAGTTTGGAGGAGAACACTACTGCTTCGGCC 1594
Db 153 TGCACCCAGCTTTGGGAGCAGTGCCCAACAAGTTTGGAGGAGAACACTACTGCTTCGGCC 212
QY 1595 GGACCGTGGAGACACAGCTGCTCGCTTTGGCAAGACCATCATGGAGGAGCAGCTGGTAC 1654
Db 213 GGACCGTGGAGACACAGCTGCTCGCTTTGGCAAGACCATCATGGAGGAGCAGCTGGTAC 272
QY 1655 TGAAGCGGTGGGCAACATCCTCATCAACCTGTATGGCATGACGCGCGTGTCTGCGGG 1714
Db 273 TGAAGCGGTGGGCAACATCCTCATCAACCTGTATGGCATGACGCGCGTGTCTGCGGG 332
QY 1715 CGAGCGCTCCATCCGATTGGGCTCCGCAACACACAGGAGTTCTCTTGGCCAACA 1774
Db 333 CGAGCGCTCCATCCGATTGGGCTCCGCAACACACAGGAGTTCTCTTGGCCAACA 392
QY 1775 CTTCTCGCTGGGAGCTTACTTGCAGAACTCTTTCAGCCTCTCTCAGCTGGACAAGTATG 1834
Db 393 CTTCTCGCTGGGAGCTTACTTGCAGAACTCTTTCAGCCTCTCTCAGCTGGACAAGTATG 452
QY 1835 CTCAGAAAACCTAGATGACGAGATTAAGAAAGTGTCT 1871
Db 453 CTCAGAAAACCTAGATGACGAGATTAAGAAAGTGTCT 489
```

RESULT 7

```
US-09-918-995-23707
; Sequence 23707, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23707
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-23707
```

Query Match 18.4%; Score 451.8; DB 11; Length 483;
Best Local Similarity 99.6%; Pred. No. 5.2e-130;
Matches 453; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1415 CCAGATCCATGAGCTTAACAGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCCGGA 1474
Db 29 CCAGATCCATGAGCTTAACAGGCCAAAGTGAGCACAGTCATGGATACCGTTGGCCGGA 88
QY 1475 GGCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGCGTGACAGGCAACCATGGAGTTG 1534
Db 89 GGCTTCGGGACTCCCTGGGCGGAACCTGTGGACCTGGGCGTGACAGGCAACCATGGAGTTG 148
QY 1535 TGCACCCAGCTTTGGGAGCAGTGCCCAACAAGTTTGGAGGAGAACACCTACTGCTTCGGCC 1594
Db 149 TGCACCCAGCTTTGGGAGCAGTGCCCAACAAGTTTGGAGGAGAACACCTACTGCTTCGGCC 208
```

```
QY 1595 GGACCGTGGAGACACTGCTCCGCTTTGGCAAGACCATCATGGAGGAGCAGCTGGTAC 1654
Db 209 GGACCGTGGAGACACTGCTCCGCTTTGGCAAGACCATCATGGAGGAGCAGCTGGTAC 268
QY 1655 TGAAGCGGTGGGCGGAACATCCCTCATCAACCTGTATGGCATGACGCGCGTGTCTGCGGG 1714
Db 269 TGAAGCGGTGGGCGGAACATCCCTCATCAACCTGTATGGCATGACGCGCGTGTCTGCGGG 328
QY 1715 CGAGCGCTCCATCCGCTTTGGCAAGACCATCATGGAGGAGCAGCTGGTAC 1774
Db 329 CGAGCGCTCCATCCGCTTTGGCAAGACCATCATGGAGGAGCAGCTGGTAC 388
QY 1775 CTTCTCGCTGGGAGCAGTTACTTGCAGAACTCTCTCAGCCTCTCTCAGCTGGACAAGTATG 1834
Db 389 CTTCTCGCTGGGAGCAGTTACTTGCAGAACTCTCTCAGCCTCTCTCAGCTGGACAAGTATG 448
QY 1835 CTCAGAAAACCTAGATGACGAGATTAAGAAAGTGTCT 1869
Db 449 CTCAGAAAACCTAGATGACGAGATTAAGAAAGTGTCT 483
```

RESULT 8

```
US-09-918-995-11034
; Sequence 11034, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11034
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-11034
```

Query Match 15.9%; Score 390.6; DB 11; Length 451;
Best Local Similarity 97.8%; Pred. No. 6.3e-111;
Matches 396; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
QY 651 AGCCACACTAAGTGAAGACAGCAAGAGCACTACATCTCAATGGCTTCCAAAGTCTGGATTAC 710
Db 47 AGCCACACTAAGTGAAGACAGCAAGAGCACTACATCTCAATGGCTTCCAAAGTCTGGATTAC 106
QY 711 TAATGGAGGACTGGCCCAATATTTTACTGTGTTTGGCAAGACTGAGTCTGTGATCTGA 770
Db 107 TAATGGAGGACTGGCCCAATATTTTACTGTGTTTGGCAAGACTGAGTCTGTGATCTGA 166
QY 771 TGGATCAGTGAAGACAAAAATCACAGCATTCATAGTAGAAGAGACTTTTGGTGGAGTCAC 830
Db 167 TGGATCATTGAAGACAAAAATCACAGCATTCATAGTAGAAGAGACTTTTGGTGGAGTCAC 226
QY 831 TAATGGAAACCCGGAAGATAAATTAAGCATTCGGGGCTCCCAACACTTGTGAAGTCCATTT 890
Db 227 TAATGGAAACCCGGAAGATAAATTAAGCATTCGGGGCTCCCAACACTTGTGAAGTCCATTT 286
QY 891 TGAACACCAAGATACCTGTGGAACACATCTTGGAGAGTCCGAGATGGGTTTAAGGT 950
Db 287 TGAATACACCAAGATACCTGTGGAACACATCTTGGAGAGCGCGGAGATGGGTTTAAGGT 346
QY 951 GGCCATGAACATCCTCAACAGCGCGGTTTTCAGCATGGGCGAGCGTCTGGCTGGGCTGCT 1010
Db 347 GGTCATGAACATCCTCAACAGCGCGGTTTTCAGCATGGGCGAGCGTCTTGGCTGGGCTGCT 406
QY 1011 CAAGAGATTGATTAATGACTGCTGAGTACGCTGACAGGAA 1055
Db 407 CAAGAGATTGATTAATGACTGCTGAGTACGCTGACAGGAA 451
```

RESULT 9

US-09-917-800A-1584
; Sequence 1584, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917, 800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 1584
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012891
US-09-917-800A-1584

Query Match 15.8%; Score 388.6; DB 10; Length 2117;
Best Local Similarity 54.5%; Pred. No. 7.5e-110;
Matches 891; Conservative 0; Mismatches 729; Indels 15; Gaps 5;

QY	179	CTTTGCCCAAGACGTTTTCCTAGGCAAAATCAAGAAAGAAAGTTCCTCCATTTCAG	238
DB	236	CTTTTGCTGTGGGAATGTTCAAGAGCCAGCTTACCCAGCAGGTGTTCCCATACCCAT	295
QY	239	AAGTTAGCAAGATGAACCTTAATGAA---ATCAATCAGTTCTTGGGACCCCTGGGAAAT	295
DB	296	CTGTGCTCAATGAAGGACAGACACAAATTTCTCAAGAGCTGTTGGACCACTGCCCCGT	355
QY	296	TOTTCACATGAAGAGTGGACTCCCGAAAAATTTGACAGGAAGAAATCCCATGAA	355
DB	356	TCCTTGAAGAATGAATGACCTGCCAAGATGACTCTTGGAGAAGTGGAGGAGCA	415
QY	356	CTTTGGAGAAATGAAGAGCTAGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTG	415
DB	416	CTTTGAGGAGCTCAAGAACTTGGGGGCAATTTGGTCTCAAGTACCCAGGAGTGGGTG	475
QY	416	GCTCGGGTCTTCCCAACCACTGTACTCAAGACTAGGGAGATCATCAGCATGATG---	472
DB	476	GTITGGGCTCTCTAATACCAGTAGCTCGCTTGGCAGAGATTGTTGGCAGTGCATGACC	535
QY	473	GCTCCATCACTGTGACCTGGCAGCGCACCGCTATTGGCTCAAGGGGATCATCTGG	532
DB	536	TTGGGTGTTAGCGTTACCTGGGAGCCCATCAGAGCATCGGTTTCAAGGCACTCTGTCT	595

QY	533	CTGGCACTGAGAGCAGAAAGCCAAATACTTTCCTAACTGGCTGCCGGGAGCACATTG	592
DB	596	ATGGCACAAGGCCAGAGGAAAAATACCTCCCGAGAGTGGCATCCCGAGCGGTTGG	655
QY	593	CAGCCTTCTGCTCAGGAGCCAGCAGTGGGAGGATGACAGCTCAATCCGGAGCAGAG	652
DB	656	CGGCTTCTGCTGACTGAGCCCTCGAGGGTCCGNTGTGGCCTCTATCCGAAGCTCAG	715
QY	653	CCACACTAAGTGAAGACAAAGACACTACATCCTCAATGGTCTCAAGGTCCTGGATTACTA	712
DB	716	CTGTACCTAGCCCTGTGGAAAGTATTATCTCTCAACGGAAGCAAGATCTGGATCAGTA	775
QY	713	ATGGAGGACTGCCCAATATTTTACTGTGTTTGCAGAGACTGAGCTCGTTGATTCTG---	769
DB	776	ATGGGGTCTGGCAGACATTTTCACTGTCTTGCACAAACGCAATTAAGATGAGGCCA	835
QY	770	ATGGATCACTGAAAGACAAATACACAGCATTCATAGTAGAAGAGACTTTGTGGAGTCA	829
DB	836	CGGGGCCGTGAAGAGAGATCACAGCTTTCGTAGTGAACGGAGCTTTGAGGGGTTA	895
QY	830	CTAATGGGAAACCCGAAAGATAAATTAGGCATTCGGGGCTCCAAACACTTTGTGAAGTCCAT	889
DB	896	CCCATGGCTCCCGAAAAAGATGGGCATCAAGGCATCTAACACATCAGAGGTGTACT	955
QY	890	TTGAAACACACAGATACCTGTGGAAACATCTTGGAGAGTTCGGAGATGGTTTAAGG	949
DB	956	TTGATGGAGTCAAGGTGCCAGCAGAGAATGCTGATGAGAGTGGGAGATGCTTCAAGG	1015
QY	950	TGGCCATGAACATCCTCAACAGCGCGGTTTCAGCATGGGCGAGCTGCTGGCTGGGCTGC	1009
DB	1016	TTGCTGTCAACATCCTCAACACGGAAGATTTGGATGGCTGCAACCTTAGCAGGCCA	1075
QY	1010	TCAAGAGATTGATGAAATGACTGCTGAGTACGCTGCAACAGGAACAGTTTAAACA	1069
DB	1076	TGAAAGCCATCATTTGCCAAGCGGTTGATCTACTACTACCGTACCCAGTTTGGGGACA	1135
QY	1070	GGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCTGATGGCTCAGAGGCTTAGC	1129
DB	1136	AAATTCACAACTTTGGGGTATCCAGAAAGCTGCTCGGATGGCTATTCGACGATG	1195
QY	1130	TCATGAGAGTATGACTACCTCACAGCAGGAGTGTGGACCAACCTGGCTTCCCGACT	1189
DB	1196	TGACTGAGTCCATGGCTTACATGCTGAGTGCCCAACATGGACCCAG---GGATCAAAAGCT	1252
QY	1190	GTCTCCAGCAGCAGCAGCTGTTCAAGTGTTCAGCTCCGAGGCGCTGCGAGTGTGTA	1249
DB	1253	TCAGATAGAGCTGCCATCAGCAAAATCTTGGCTCGGAGGCGGCTGGAAGTGCAG	1312
QY	1250	GTGAGCGCTGACAGATCCTCGGGGCTTGGGCTACACAAAGGAGTATCCGTAGCAGCA	1309
DB	1313	ATGAGTGCATCCAGATAATGGGGGCTTTCATGAAGGAACCAAGGAGTATGAGCGTG	1372
QY	1310	TACTGGTGCACCGCATCTCTCATCTTCAGGGAACCAATGAGATTCTCCGATGT	1369
DB	1373	TGCTCCGAGATATTCGAATCTCCGATCTTTGAAGGGAACAAATGACATCTTCGACTGT	1432
QY	1370	ACATGCGCTGACGGGTCTCAGCATGCGCGCGCATCTCTGACTACCAAGTATCCATGAGC	1429
DB	1433	TTGTGGCTCTACAGGCTGCATGGACAAAGAAAGGAACATCAGGGACTTGGTAATGCC	1492
QY	1430	TTAAACAGGCAAGTGAAGCAGTATGATACCGTTGGCGGAGGCTTCGGGACTCC	1489
DB	1493	TAAAGATCTCTTGGAAATGTTGGCTCTCATAGGAGAAAGCAACAGCTGAGGC	1552
QY	1490	TGGCCCACTGTGGACCTGGCTGACAGCAACCATGAGTTGTGACCCCGCTCTTG	1549
DB	1553	GGCGAGCAGGATTTGGCAGTGGTCTGA---GTCTCTCGGAATTTCCACCCAGAGTTGA	1609
QY	1550	CGGACAGTGGCAACAAGTTTTCAGGAGAACACCTACTCTTCGCGCGGACCGTGGAGAC	1609
DB	1610	GTCCAGTGGTGAACCTGCAGTGCAGCTCTGGAACAATTTGCCACTGTAGTGGAGCGA	1669
QY	1610	TGCTGCTCGGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGCAAGCGGTGCCA	1669

Db 1670 ACCTCATGACGACAAAGGATGTGCAATGACAGTTCTCTGCGAGCGATGGCAG 1729
QY 1670 ACATCTCTATCAACCTGTATGGCATGACGGCGGTGCTGTGCGGGCCGACCGCTCCATCC 1729
Db 1730 ATGGAGCCATTGACCTCTACGCCATGGTGGTGTCTCTCCAGAGCCTCAAGATCCCTGA 1789
QY 1730 GCATTGGGCTCCGCAACACCAACGAGGTTCTCTTGGCCAAACACTTCTGGTGAAG 1789
Db 1790 GTGAGGCTACCCGACAGCAGCATGAGAAATGCTCTGTGTAGTTGGTGCATTGAGG 1849
QY 1790 CTTACTTCCGAATC 1804
Db 1850 CTGCAACACGGATTC 1864

RESULT 10

US-09-918-995-16927
; Sequence 16927, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16927
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-16927

Query Match 10.7%; Score 261.4; DB 11; Length 490;
Best Local Similarity 99.3%; Pred. No. 1.4e-70;
Matches 273; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1626 CAGACCATCATGGAGGAGCAGTGTACTGAAGCGGTGGCCAACTCCTCATCAACT 1685
Db 217 CCAGACCATCATGGAGGAGCAGTGTACTGAAGCGGTGGCCAACTCCTCATCAACT 276
QY 1686 GTATGGCATGACGGCGTCTGTCGGGCCAGCGCTCCATCCGATGGGCTCCGCAA 1745
Db 277 GTATGGCATGACGGCGTCTGTCGGGCCAGCGCTCCATCCGATGGGCTCCGCAA 336
QY 1746 CCAGACCATGAGGTCTCTTGGCCAACTCTGCTGGGAAGCTTACTTGCAGAACT 1805
Db 337 CCAGACCATGAGGTCTCTTGGCCAACTCTGCTGGGAAGCTTACTTGCAGAACT 396
QY 1806 CTTGACCTCTCTGAGTGGACAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAGAA 1865
Db 397 CTTGACCTCTCTGAGTGGACAGTATGCTCCAGAAAACCTA-ATGAGCAGATTAAGAA 455
QY 1866 ACTGTCCAGCAGATCTTGTGAGAACGAGCCTATA 1900
Db 456 ACTGTCCAGCAGATCTTGTGAGAACGAGCCTATA 490

RESULT 11

US-09-822-849A-480
; Sequence 480, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary

; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakr
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 480
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-480

Query Match 10.5%; Score 258.6; DB 9; Length 312;
Best Local Similarity 96.7%; Pred. No. 7.8e-70;
Matches 264; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 23 TAAGAAGGGGAGACTGAGGCTGGGGAACATCGGCAGCATGAGCGGCTGCGGGC 82
Db 13 TAAGAAGGGGAGACTGAGGCTGGGGAACATCGGCAGCATGAGCGGCTGCGGGC 72
QY 83 TCTTCTCGGCACACGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCTCTACCGCGA 142
Db 73 TCTTCTCGGCACACGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCTCTACCGCGA 132
QY 143 ACCGGCGCTACTGCGCACACCGCGCTGTACGAGCTTCCGCAAGAGCTTTTCCTAG 202
Db 133 ACCGGCGCTACTGCGCACACCGCGCTGTACGAGCTTCCGCAAGAGCTTTTCCTAG 192
QY 203 GCAAAATCAAGAAGAAAGTTCCTCCATTTCCAGAGTTAGCCAAAGATGAACCTTAATG 262
Db 193 GCAAAATCAAGAAGAAAGTTCCTCCATTTCCAGAGTTAGCCAAAGATGAACCTTAATG 252
QY 263 AAATCAATCAGTTCTTGGGACCCGCGGAAAAAT 295
Db 253 AAATCAATCAGTTCTTGGCTAACTTCTGGAAT 285

RESULT 12

US-09-918-995-4541
; Sequence 4541, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4541
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-4541

Query Match 7.6%; Score 187.4; DB 11; Length 402;
Best Local Similarity 94.6%; Pred. No. 1.5e-47;
Matches 194; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 11 TCCCTCGGCGCTAAGAAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGCAGCATGA 70
|||||

Db	1	TCCTCCGGCGCTTAAGAGGGAGACTGAGGCTGTGGTGTGGAACATCTGCCACATCA	60
QY	71	GCGGCTGCGGGCTTCTCTCCGCGCACACACCGCTGCGGCTGCTCCCTGCCGGGTCTGGTGG	130
Db	61	GCGGCTGCGGGCTTCTCTCCGCGCACACACCGCTGCGGCTGCTCCCTGCCGGGTCTGGTGG	120
QY	131	TCCTACCGGGAACCGGGGCTTACTGCGCACACCGCGCTCTACGAGCTTTCGCCAAG	190
Db	121	TATCTACCGCAACCGGGGCTTACTGCGCACCATCCGCTCTACGAGCTTTCGCCAAG	180
QY	191	AGCTTTTCCTAGGCAAAATCAAGAA	215
Db	181	AGCTTTTCCTAGGCAAAATCAAGAA	205
RESULT 13			
US-09-796-692-8175			
; Sequence 8175, Application US/09796692			
; Publication No. US20020198362A1			
; GENERAL INFORMATION:			
; APPLICANT: Gaiger, Alexander			
; APPLICANT: Mannion, Jane			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY			
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES			
; FILE REFERENCE: 2077.001200			
; CURRENT APPLICATION NUMBER: US/09/796,692			
; CURRENT FILING DATE: 2001-03-01			
; PRIOR APPLICATION NUMBER: 60/186,126			
; PRIOR FILING DATE: 2000-03-01			
; PRIOR APPLICATION NUMBER: 60/190,479			
; PRIOR FILING DATE: 2000-03-17			
; PRIOR APPLICATION NUMBER: 60/200,545			
; PRIOR FILING DATE: 2000-04-27			
; PRIOR APPLICATION NUMBER: 60/200,303			
; PRIOR FILING DATE: 2000-04-28			
; PRIOR APPLICATION NUMBER: 60/200,779			
; PRIOR FILING DATE: 2000-04-28			
; PRIOR APPLICATION NUMBER: 60/200,999			
; PRIOR FILING DATE: 2000-05-01			
; PRIOR APPLICATION NUMBER: 60/202,084			
; PRIOR FILING DATE: 2000-05-04			
; PRIOR APPLICATION NUMBER: 60/206,201			
; PRIOR FILING DATE: 2000-05-22			
; PRIOR APPLICATION NUMBER: 60/218,950			
; PRIOR FILING DATE: 2000-07-14			
; PRIOR APPLICATION NUMBER: 60/222,903			
; PRIOR FILING DATE: 2000-08-03			
; PRIOR APPLICATION NUMBER: 60/223,378			
; PRIOR FILING DATE: 2000-08-07			
; NUMBER OF SEQ ID NOS: 9597			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 8175			
; LENGTH: 568			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-796-692-8175			
Query Match 6.7%; Score 165.2; DB 10; Length 568;			
Best Local Similarity 58.0%; Pred. No. 1.6e-40;			
Matches 312; Conservative 0; Mismatches 223; Indels 3; Gaps 1;			
QY	848	ATAAATAGGCATTCGGGCTCCACACTTGTGAAGTCCATTTTGAACACACACAGATAC	907
Db	3	AGAAGATGGCATCAAGCGCTTCAACACACAGCAGAGGTTCTTTGATGAGTACGGGTGC	62
QY	908	CTGTGGAACATCTTGGAGATCGGAGTGGTTTAAAGTGGCCATGAACATCCCTCA	967
Db	63	CATCGGAGAACGTGCTGGGTGAGGTTGGAGTGGCTTCAAGTTGCCATGCACATCCCTCA	122
QY	968	ACAGCGCGCGGTTACAGATCGGACGCGTGGTGGCTGCTCAAGAGATTGATTGAA	1027

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 23:10:50 ; Search time 5159.14 Seconds
(without alignments)
11551.252 Million cell updates/sec

Title: US-09-945-326-1
Perfect score: 2452
Sequence: 1 cgtgtgtgtcctcgtgcgc.....actgttaaaaaaaaaaaaaa 2452

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estnu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_nam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1683.6	68.7	1914	11 BC041572	BC041572 Homo sapi
2	1382	56.4	2432	11 AK075984	AK075984 Mus muscu
3	1379.4	56.3	2446	11 AK031820	AK031820 Mus muscu
4	1082.2	44.1	2906	11 AK028571	AK028571 Mus muscu

5	1026	41.8	1123	11 AF078854	AF078854 Homo sapi
6	952.8	38.9	1071	12 BM080820	BM080820 AGENCOURT
7	940.4	38.4	1201	9 AL539220	AL539220 AL539220
8	937.8	38.2	1201	13 BX458691	BX458691 BX458691
9	879.2	35.9	1201	9 AL567735	AL567735 AL567735
10	869.6	35.5	969	9 AL549834	AL549834 AL549834
11	867.2	35.4	1010	12 BM552756	BM552756 AGENCOURT
12	859.4	35.0	1079	12 BM561207	BM561207 AGENCOURT
13	854	34.8	1206	9 AL582366	AL582366 AL582366
14	852.4	34.8	890	13 BQ938039	BQ938039 AGENCOURT
15	852.4	34.8	1201	9 AL560956	AL560956 AL560956
16	847.2	34.6	905	13 BQ541780	BQ541780 AGENCOURT
17	844.4	34.4	1020	13 BUI96321	BUI96321 AGENCOURT
18	838.6	34.2	1186	9 AL578429	AL578429 AGENCOURT
19	833.2	34.0	914	9 AL524997	AL524997 AL524997
20	822.2	33.5	1023	13 BQ072180	BQ072180 AGENCOURT
21	813.4	33.2	851	13 BUI76038	BUI76038 AGENCOURT
22	810.6	33.1	884	14 CA488487	CA488487 AGENCOURT
23	810	33.0	922	13 BQ528453	BQ528453 AGENCOURT
24	809.6	33.0	1199	9 AL556077	AL556077 AL556077
25	808	33.0	831	12 B1524078	B1524078 603052267
26	807.8	32.9	1023	12 BM09053	BM09053 AGENCOURT
27	805.2	32.8	875	13 BQ882994	BQ882994 AGENCOURT
28	799.6	32.6	885	13 BUI91988	BUI91988 AGENCOURT
29	798.8	32.6	888	13 BQ960770	BQ960770 AGENCOURT
30	798.2	32.6	848	9 AL557534	AL557534 AL557534
31	797.6	32.5	937	13 BQ279135	BQ279135 AGENCOURT
32	793.2	32.3	978	12 B1762564	B1762564 603048696
33	792	32.3	980	12 BQ054246	BQ054246 AGENCOURT
34	790.4	32.2	857	9 AL557535	AL557535 AL557535
35	788.4	32.2	867	10 BG674212	BG674212 602619872
36	784.4	32.0	799	12 B1914477	B1914477 603182269
37	780.2	31.8	891	13 BX328269	BX328269 BX328269
38	772.4	31.5	851	10 BG757623	BG757623 602711206
39	767.4	31.3	885	13 BUI56414	BUI56414 AGENCOURT
40	756	30.8	905	9 AL524996	AL524996 AL524996
41	755.6	30.8	893	12 B1772602	B1772602 603053154
42	754.2	30.8	973	10 BE792118	BE792118 601581681
43	750.6	30.6	771	9 AU142885	AU142885 AU142885
44	748.8	30.5	761	12 B1821212	B1821212 603034922
45	747.4	30.5	784	12 B1829609	B1829609 603079347

ALIGNMENTS

RESULT 1	BC041572	Homo sapiens, Similar to NP0002 protein, clone IMAGE:3952086, mRNA.	1914 bp	mRNA	linear	HTC 02-JAN-2003
LOCUS	BC041572					
DEFINITION	BC041572					
ACCESSION	BC041572.1	GI:27469662				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
REMARK						
COMMENT						

BC041572 Homo sapiens, Similar to NP0002 protein, clone IMAGE:3952086, mRNA.
BC041572 Homo sapiens
BC041572.1 GI:27469662
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapsb@remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgen@nri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 44 Row: c Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis

This clone has the following problem: frame shifted.

FEATURES

source	1. 1914	Location/Qualifiers		
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:3952086"			
	/tissue_type="Ovary, adenocarcinoma"			
	/clone_lib="NIH_MGC_9"			
	/lab_host="DH10B-R"			
	/note="Vector: pOTB7"			
BASE COUNT	477 a 482 c 522 g 433 t			
ORIGIN				
Query Match 68.7%; Score 1683.6; DB 11; Length 1914;				
Best Local Similarity 97.7%; Pred. No. 0;				
Matches 1741; Conservative 0; Mismatches 9; Indels 32; Gaps 2;				
QY	698 AGGCTGGATTACTAATGGAGGACTGGCCCAATATTTTACTGTGTTGCAAGACTGAGG	757		
DB	133 AGGCTGGATTACTAATGGAGGACTGGCCCAATATTTTACTGTGTTGCAAGACTGAGG	192		
QY	758 TCGTTGATTCGATGGATCAGTGAAAGACAAATATCAGCATTCATAGTAGAAGAGACT	817		
DB	193 TCGTTGATTCGATGGATCAGTGAAAGACAAATATCAGCATTCATAGTAGAAGAGACT	252		
QY	818 TTGTTGGAGTCACTAATGGGAAACCCGAAGATAAATAGGATTCGGGGCTCCCAACTT	877		
DB	253 TTGTTGGAGTCACTAATGGGAAACCCGAAGATAAATAGGATTCGGGGCTCCCAACTT	312		
QY	878 GTGAAGTCCATTGAAACACCAAGATACCTGTGGAAACATCCTTGGAGAGTCGGAG	937		
DB	313 GTGAAGTCCATTGAAACACCAAGATACCTGTGGAAACATCCTTGGAGAGTCGGAG	372		
QY	938 ATGGGTTTAAAGTGGCCATGAACATCCCTCAACAGCGCGGTTTCAGCATGGGAGCTCG	997		
DB	373 ATGGGTTTAAAGTGGCCATGAACATCCCTCAACAGCGCGGTTTCAGCATGGGAGCTCG	432		
QY	998 TGGCTGGGCTGCTCAAGAGATGATT-----GAAATGACTGCTGAGTACGCTGCACAAAGG	1053		
DB	433 TGGCTGGGCTGCTCAAGAGATGATTGTTGAAATGACTGCTGAGTACGCTGCACAAAGG	492		
QY	1054 AAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTGATT-----	1092		
DB	493 AAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTGATTTCCTTTGTTTCTGTTTCTG	552		
QY	1093 -----CAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATGATGAC	1145		
DB	553 AAACACTCCAGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGGAGATGATGAC	612		
QY	1146 CTACCTCACAGCAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGC	1205		
DB	613 CTACCTCACAGCAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGC	672		
QY	1206 CATGTTGAAGGTTTCAGTCCGAGCGCCCTGGCAGTGTGTGAGTGGAGGCTGCAGAT	1265		

Db	673 CATGGTGAAGGTGTTTACGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGCGCTGCAGAT	732
QY	1266 CTTGGGGGCTTGGGCTACACAAGGGACTATCCGTACGAGCGCATCTGGTGCACACCCG	1325
DB	733 CTTGGGGGCTTGGGCTACACAAGGGACTATCCGTACGAGCGCATCTGGTGCACACCCG	792
QY	1326 CATCCTCCTCATCTTCGAGGAAACCAATGAGATTCCTCCGATGTACATCCCTGACGGG	1385
DB	793 CATCCTCCTCATCTTCGAGGAAACCAATGAGATTCCTCCGATGTACATCCCTGACGGG	852
QY	1386 TCTCAGCATGCGCGCGCATCTCTGACTACCAAGGATCCATGAGCTTAAACAGGCCAAAGT	1445
DB	853 TCTCAGCATGCGCGCGCATCTCTGACTACCAAGGATCCATGAGCTTAAACAGGCCAAAGT	912
QY	1446 GAGCACAGTCAATGATACCGTTGGCGGAGGCTTGGGACTCCCTGGCGGCAACTGTGGA	1505
DB	913 GAGCACAGTCAATGATACCGTTGGCGGAGGCTTGGGACTCCCTGGCGGCAACTGTGGA	972
QY	1506 CTTGGGCTGACAGCAACCATGAGTTGTGACCCCGAGTCTTGGGAGTGTGCGGACAA	1565
DB	973 CTTGGGCTGACAGCAACCATGAGTTGTGACCCCTAGTCTTGGGAGTGTGCGGACAA	1032
QY	1566 GTTTGAGGAGAACCTTACTGCTTGGCGGAGCGGTGGAGACACTGTGCTCCGCTTTGG	1625
DB	1033 GTTTGAGGAGAACCTTACTGCTTGGCGGAGCGGTGGAGACACTGTGCTCCGCTTTGG	1092
QY	1626 CAAGACCATCATGAGGAGCAGCTGTTGAGCGGCTGAGCGGCTGAGCACTCTCATCACT	1685
DB	1093 CAAGACCATCATGAGGAGCAGCTGTTGAGCGGCTGAGCGGCTGAGCACTCTCATCACT	1152
QY	1686 GTATGGCATGACGCGCGTGTCTGCGGCGGAGCGGCTCCATCCGCAATGGGCTCCGAA	1745
DB	1153 GTATGGCATGACGCGCGTGTCTGCGGCGGAGCGGCTCCATCCGCAATGGGCTCCGAA	1212
QY	1746 CCACGACCAAGAGTTCCTTGGCGCAACACTTCTGCGTGGAGGCTTACTTGCAGAACT	1805
DB	1213 CCACGACCAAGAGTTCCTTGGCGCAACACTTCTGCGTGGAGGCTTACTTGCAGAACT	1272
QY	1806 CTTGAGCTCTCTCAGCTGGACAACTATGCTCCAGAAACCTTAGATGAGCAGATTAAGAA	1865
DB	1273 CTTGAGCTCTCTCAGCTGGACAACTATGCTCCAGAAACCTTAGATGAGCAGATTAAGAA	1332
QY	1866 AGTGTCCCGAGCAGATCCTTGGAGAGCGCTTATCTGTGCGCCACCTCTGGACAGGAC	1925
DB	1333 AGTGTCCCGAGCAGATCCTTGGAGAGCGCTTATCTGTGCGCCACCTCTGGACAGGAC	1392
QY	1926 ATGCTCAGCAGGAGGACGTGTCCTGCTACCGCGCGGCTTACCCATGGCCCGTGTGT	1985
DB	1393 ATGCTCAGCAGGAGGACGTGTCCTGCTACCGCGCGGCTTACCCATGGCCCGTGTGT	1452
QY	1986 GGATGACTGTTACTCTTTTTCAGAAAGTGTGGGATTTATCAGGTTTAAAGCTTTTGT	2045
DB	1453 GGATGACTGTTACTCTTTTTCAGAAAGTGTGGGATTTATCAGGTTTAAAGCTTTTGT	1512
QY	2046 CCCCCTGTCGACCTGAAGGTTGTCCTGCTGGCTGGAGAGGCTCTTCCAGAGTTTGTACC	2105
DB	1513 CCCCCTGTCGACCTGAAGGTTGTCCTGCTGGCTGGAGAGGCTCTTCCAGAGTTTGTACC	1572
QY	2106 TGCAGGAGTGTCTCTTAACAGGACCATCACAGCTTCTGAATGAGCGGAGAGAGAGAA	2165
DB	1573 TGCAGGAGTGTCTCTTAACAGGACCATCACAGCTTCTGAATGAGCGGAGAGAGAGAA	1632
QY	2166 TGGAAATGCTGACCCCTGGAACTGGCGGGTATTTCTGTCATTCAGGAGACACCATAGTG	2225
DB	1633 TGGAAATGCTGACCCCTGGAACTGGCGGGTATTTCTGTCATTCAGGAGACACCATAGTG	1692
QY	2226 AAACCTGGGCTTATGCTGCTGCCCTCCAGGCTGTGAGGTGGGTGGGACCTGTGTCAAGTG	2285
DB	1693 AAACCTGGGCTTATGCTGCTGCCCTCCAGGCTGTGAGGTGGGTGGGACCTGTGTCAAGTG	1752
QY	2286 TGGATAGCCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTCAAGAGCTGTCTGT	2345
DB	1753 TGGATAGCCATTTCTGCTCAACACACATTTCTTAAGAAACAGCTTCAAGAGCTGTCTGT	1812

```

QY 2346 GGTCAATTCATTAACTAGACAGCAGCCACTTAAACATGTCACCAAGAACCATTTAAACA 2405
      |||||||
Db 1813 GGTCAATTCATTAACTAGACAGCAGCCACTTAAACATGTCACCAAGAACCATTTAAACA 1872
      |||||||
QY 2406 AAGAAATATAAATGTCACAACTCTGTGACTGCTTAAAAAATAA 2447
      |||||||
Db 1873 AAGAAATATAAATGTCACAACTCTGAAAAAATAAATAAATAA 1914
      |||||||

RESULT 2
ACCESSION AK075984
LOCUS Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
DEFINITION enriched library, clone:2600017P15 product:VERY-LONG-CHAIN ACYL-COA
DEHYDROGENASE VLCAD homolog (Homo sapiens), full insert sequence.
ACCESSION AK075984
VERSION 1
KEYWORDS GI:26344893
SOURCE HTC; CAP trapper.
ORGANISM Mus musculus (house mouse)

REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipeptide sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
REFERENCE 4
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staib, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660

```

PUBMED
REFERENCE
AUTHORS

TITLE

JOURNAL

REFERENCE
AUTHORSTITLE
JOURNAL

COMMENT

FEATURES
SOURCE

CDS

BASE COUNT
ORIGIN

Query Match

11217851
5The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 2432)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Izawa, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyezaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toyama, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1..2432

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM,DB:2600017P15"

/db_xref="MGI:1895339"

/db_xref="taxon:10090"

/clone="2600017P15"

/tissue_type="whole body"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10 days embryo"

29..1906

/note="unnamed protein product; VERY-LONG-CHAIN ACYL-COA
DEHYDROGENASE VLCAD homolog (Homo sapiens) (SPTR1AAL56011,
evidence: FASTV, 85.6%id, 99.6%length, match=1869)
putative"

/codon_start=1

/protein_id="BAC36096.1"

/db_xref="GI:26344894"

/db_xref="MGI:1914272"

/translation="MSGCVLSRGATAAAARASRVLRREFARRRPLHTLSQCSFA
KELFLGNEQKGVFPFVPSQHEISEINQFVPLEKFEETVEDSRKIDQEGIPVDTL
EKLSLGLGIVQPEEYGGLSNTMYARGLGEIISLDAITVTLAAHQAIGLKIILV
ENKEQAKYLPKLSGEHIAAFCLETPASGSDAASIQTRATLSDEKQYFILINGSKVNI
TNGLIANTVFVAKTEVVDSGSKDKMTAFIVERDFGTTNGKPEOKLIGRSGTCE
VFENRTFVENVLGVGSGFKVANNILNSGRFSNGSVAGNKKLLIELTAYACTRRK
QFNRLSEGLIOEKFAALMAQKAYMESMAYLTSCMLDQPGFPCDSEAMVAFVSE
AAWQYSEALQILGSGFMKDIPTERMURDARILLIFEGTNEVRLFTALTGLQHGR
ILTSRIKESLGNVTVMETGRKRLSLGRTDGLTGLGVHPHSRSGSKLEND
HYFGRTVETLLRFGKNIVEEQLVRANILNLYGMTAVLSRASRSTRIGLRNHD
HEVLIANNFCVEAYFQNLFSLSQDKNAPENLDQIKKVSRIQLEKRAYICAPLDR
S"

640 a 564 c 652 g 576 t

56.4% Score 1382; DB 11; Length 2432;

Best Local Similarity 84.3%; Pred. No. 1.16e-292;		Matches 1556; Conservative 0; Mismatches 290; Indels 0; Gaps 0;	
QY	98	CGGCTCGGCTCGTGCCTGCGGGTCTGGTGGTCTCTACCGCGAACCAGCGGGTCTACTGC	157
Db	72	CGGCTCGGCGGGAGCTTCCCGGGTCTGAGGGAATTCACGCGAGGCGCGCGCTGC	131
QY	158	GCACCAAGCCGCTGTACGAGCTTTGCGCAAGAGCTTTTCTAGGCAAAATCAAGAGA	217
Db	132	ACACCAAGCCTCGAGAGCTGCTGGTTCGCAAGAGCTCTTCTGCGCAACATCGAGCAGA	191
QY	218	AGAAGCTTTCCGATTTCCAGAGCTTAGCCAGATGAATTAATCAATCAATCAATCTCT	277
Db	192	AAGGAGTTTCCGATTTCCAGAGCTTAGCCAGATGAATTAATCAATCAATCAATCTCT	251
QY	278	TGGGACCCGTGGAAAAATCTTCACTGAAGAGGTGGACTCCGCAAAAAATTGACAGGAG	337
Db	252	TGGGACCCGTGGAAAAATCTTCACTGAAGAGGTGGACTCCGCAAAAAATTGACAGGAG	311
QY	338	GGAAATCCGAGATGAATTTGGGAAATTTGAAGAGCTTAGGCTTTTGGGCTGCAAG	397
Db	312	GAATAATCCGAGATGAATTTGGGAAATTTGAAGAGCTTAGGCTTTTGGGCTGCAAG	371
QY	398	TCCAGAGAAATATGGTGGCTGGGCTTCTCAACACCATGTACTCAAGACTAGGGAGA	457
Db	372	TCCAGAGAAATATGGTGGCTGGGCTTCTCAACACCATGTACTCAAGACTAGGGAGA	431
QY	458	TCATCAGCATGGATGGGTCATCATCTGTGACCTTGGCAGCGCACAGGCTATTGGCTCA	517
Db	432	TCATCAGCATGGATGGGTCATCATCTGTGACCTTGGCAGCGCACAGGCTATTGGCTCA	491
QY	518	AGGGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCAAAATCTTGCCTAAATCGCGT	577
Db	492	AGGGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCAAAATCTTGCCTAAATCGCGT	551
QY	578	CCGGGAGCAATTCAGAGCTTTCGCTCAGGAGCAGCCAGTGGGAGCATGAGCCT	637
Db	552	CTGGGAAACATCGCAGCTTCTGCTTAACAGAGCAGCCAGTGGGAGCATGAGCCT	611
QY	638	CAATCCGAGCAGAGCAGCAGTAACTGAAGCAGAAAGCAATCATCTCAATGGCTCCA	697
Db	612	CAATCCGAGCAGAGCAGCAGTAACTGAAGCAGAAAGCAATCATCTCAATGGCTCCA	671
QY	698	AGGCTGGATTAATAGGAGCTGGCCAAATATTTTACTGTGTGTTGCAAGACTGAGG	757
Db	672	AGGTTGGATCACCAGTGGAGCTGGCCAAATATTTTACTGTGTGTTGCAAAACCCAGG	731
QY	758	TCGTTGATCTGATGGATCAGTGAAGCAGAAATACAGCATTCATAGTAGAAAGAGACT	817
Db	732	TGGTCGATCTGATGGTTGGAACAGACAAATGACCGCATTCATAGTAGAAAGAGACT	791
QY	818	TTGTTGGAGTCACTAATGGGAAACCCGAGATAAATTAGGCATTCGAGGCTCCACACTT	877
Db	792	TGGGCGGAATCACTAATGGGAAACCCGAGATAAATTAGGCATTCGAGGCTCCACACTT	851
QY	878	GTGAGTCCATTTGAAACACACAGATACCTGTGGAAACATCTTGGAGAGTCCGAG	937
Db	852	GTGAGTCCATTTGAAACACACAGATACCTGTGGAAACATCTTGGAGAGTCCGAG	911
QY	938	ATGGGTTAAAGTGGCCATGAACATCTTCAACAGCGCCGCTTCAAGATGGCAGCGTCG	997
Db	912	GGGGCTTTAAGTACCATGAACATCTTCAACAGTGGCGATTCAGCATGGCAGTGTG	971
QY	998	TGGCTGGGCTCAAGAGATGATTAATGACTGCTGAGTACGCTGACAGGAAAC	1057
Db	972	TGGCGGGATGCTGAAGAGCTGATCGAATGACTGCTGAGTATGCTGTACGAGGAAC	1031
QY	1058	AGTTAAACAGAGCTCAGTGAATTTGGATGATTCAGGAGAAATTTGCATGTAGGCTC	1117
Db	1032	AGTTAAACAGAGCTCAGTGAATTTGGATGATTCAGGAGAAATTTGCATGTAGGCTC	1091
QY	1118	AGAAGCTTACGTCATGGAGATGATGACCTACCTCAGCAGCGGATGCTGACCAACCTG	1177

Db	1092	AGAAGCATACGTAATGAGAGATGATGCTTACCTCACTTACAGGATGCTAGACCAACAG	1151
QY	1178	GCATTCCCGAGCTGCTCCATCAGGAGCAGCCATGTTGAAGTGTTCAGCTCCGAGCGCGCT	1237
Db	1152	GATTTCGCGAGCTGCTCTATCGAGCAGCCATGTTGAAGTGTTCAGCTTCGAGCTGCTC	1211
QY	1238	GGCAGTGTGTAGTGGCGCTCGAGATCTCGGGGCTTGGGCTACACAAAGGAGCTATC	1297
Db	1212	GGCAGTGTGTAGGAGGCTCTGCAGATCTTGGGGCTCAGGCTACATGAAGGAGCTACC	1271
QY	1298	CGTACGAGCGCACTGCGGTGACACCCGATCTCTCCCTCATCTTCGAGGGAACCAATGAGA	1357
Db	1272	CCTATGAGCGCATCTGCGTATGCGCGCATCTTCTTAATCTTTGAGGGAACCAATGAGA	1331
QY	1358	TTCTCCGATGATACGCGCTGACGGTCTGACAGCATGCGCGCGCATCTCTGACTACCA	1417
Db	1332	TTCTCCGATGATACGCGCTGACGGTCTGACAGCATGCGCGCGCATCTCTGACTACCA	1391
QY	1418	GGATCCATGAGCTTAAACAGCGCAAGTGGACAGCAGTCAATGATACCTTGGCGCGGAGC	1477
Db	1392	GGATCCATGAGCTTAAACAGCGCAAGTGGACAGCAGTCAATGATACCTTGGCGCGGAGC	1451
QY	1478	TTGGGAGCTCCCTGGCGGAACTGTGGACCTTGGGCTTAAACAGCGGATCTTGGTGTGTC	1537
Db	1452	TTGGGAGCTCCCTGGCGGAACTGTGGACCTTGGGCTTAAACAGCGGATCTTGGTGTGTC	1511
QY	1538	ACCCGAGCTTGGCGGAGTGGCAACAGTTTGGAGGAAACACCTACTGCTTCGCGCGGA	1597
Db	1512	ACCCGAGCTTGGCGGAGTGGCAACAGTTTGGAGGAAATGTCCATTAATCTTGGCGGGA	1571
QY	1598	CGGTGAGACACTGCTGCTCGGCTTTGGCAAGACCATCATGAGGAGCAGCTGTGACTGA	1657
Db	1572	CGGTGAGACACTGCTGCTCGGCTTTTGGAAAGCAATTTGAGAGAACAGCTGTGTGTA	1631
QY	1658	ACCGGCTGCGCAACATCTCATCACTTATGATGATGATGATGATGATGATGATGATGAT	1717
Db	1632	ACCGGCTGCGCAACATCTCATCACTTATGATGATGATGATGATGATGATGATGATGAT	1691
QY	1718	GCCTGCTCCATCGCATTTGGGCTCCGCAACAGCAGCAGGAGTCTCTTGGCCAAACCT	1777
Db	1692	GCCTGCTCCATCGCATTTGGGCTCAGGAATCATGATCAGGAGTCTCTTGGCCAAACCT	1751
QY	1778	TCCTGCTGGAAGCTTACTTGCAGATCTCTTCAGCTCTCTCAGCTGAGCAGATGCTC	1837
Db	1752	TCCTGCTGGAAGCTTACTTGCAGATCTCTTCAGCTCTCTCAGCTGAGCAGATGCTC	1811
QY	1838	CAGAAACCTGATGAGCAGATTAAGAAAGTGTCCAGCAGATCTCTTGAAGAGCGAGCT	1897
Db	1812	CAGAAACCTGATGAGCAGATTAAGAAAGTGTCTCGGAGATCTCTTGAAGAGCGAGCT	1871
QY	1898	ATATCTGTGCCCCCTCTGAGCAGACATGCTGAGGAGGAGCA	1943
Db	1872	ATATCTGTGCCCCCTCTGAGCAGACATGCTGAGGCTGGAACCA	1917

RESULT 3
AK031820
LOCUS
DEFINITION
Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
enriched library, clone:630407H09 product:VERY-LONG-CHAIN ACYL-CoA
DEHYDROGENASE VLCD homolog [Homo sapiens], full insert sequence.
AK031820
GI:26327642
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253

AK031820 2446 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
enriched library, clone:630407H09 product:VERY-LONG-CHAIN ACYL-CoA
DEHYDROGENASE VLCD homolog [Homo sapiens], full insert sequence.
AK031820
GI:26327642
VERSION
HTC; CAP trapper.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253

PUBMED
REFERENCE
AUTHORS

10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumura, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4

PUBMED
REFERENCE
AUTHORS

TITLE

10349636

5 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staibil, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kaniya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
6

TITLE

10349636

7 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
6 (bases 1 to 2446)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE

10349636

8 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
6 (bases 1 to 2446)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE

10349636

9 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
6 (bases 1 to 2446)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE

10349636

10 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
6 (bases 1 to 2446)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source
1. .2446
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6330407H09"
/db_xref="taxon:10090"
/clone="6330407H09"
/sex="male"
/tissue_type="medulla oblongata"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
42. .1919
/note="unnamed protein product: VERY-LONG-CHAIN ACYL-CoA DEHYDROGENASE VLAD homolog [Homo sapiens] (SPTRIAL56011, evidence: FASTY, 85.6%ID, 99.6%length, match=1869)
putative"
/codon_start=1
/protein_id="BAC27565.1"
/db_xref="GI:26327643"
/translation="MSGCVLLSRGATAAGAAASRVLRFTARRRPLHTSIQSCSFA
KELFLNTEQGVFPPEVSOHELSEINQFVGPLEKFFTEVDSKIDQEKIPVDTL
EKLGLGFIQVPEYGGISLNTMYARLGEIISLDASITVTLAAHQHGLGILV
GNEEKAKYLPKLSGGEHIAFCLTEPASPASDAISQTRATLSEDKYFLINGSKWI
TNGGLANITFAKTEVDSDQKMTAFIVDFGGITNGKPLDITRGSTNCE
VFENRTPVENLGEVGGGVKAMNINLSFSGSAGVAKLKLIELTAETKTRK
QPNRLSEFLQIEQFALMAQKAYVSNMAYLTSMQDPGFPDCSI EAANVKVFSSE
AAWQCSALQILGGSGYMKDYPYERMLRDARILLIPETNEILRLFTALGLOHGR
ILTSRIKELKSNVTVMETIGRLRSLDRTVDLGLDGVVHPSLGSANKLEEN
VHYFGRTVETLLRFGKNVREQLVKRVANVLNLYGTAVLSRASRSIRIGLNHD
HEVLLANNFCVEAYFQNLFSLSQLDKNAPENLDQIKKVSQIILEKRAYICAHPLDRA
S"

CDS

BASE COUNT 641 a 567 c 660 g 578 t
ORIGIN
Query Match 56.3%; Score 1379.4; DB 11; Length 2446;
Best Local Similarity 84.2%; Pred No. 5,9e-292;
Matches 1554; Conservative 0; Mismatches 291; Indels 0; Gaps 0;
QY 99 GCCTCGGCTCGTCCGCGGGTCTGGTGGTCTCTACCGCGAACCGCGGCTACTCGG 158
Db 86 GCCTCGGCGGAGCTTCCCGGGTCTGAGGGAATTCACGCGAGGCGCGCGCTGCA 145
QY 159 CACGAGCGGCTGTACGAGCTTTCGCCAAAGAGCTTTCTCTAGCGAAATCAAGAGAA 218
Db 146 CACGAGCTCGAGAGCTGCTCGTTCGCCAAGAGGCTCTTCTGCGCAACATCGAGCAGAA 205
QY 219 AGAAGTTTTCATTTCCAGAAAGTAGCAAGATGAATTAATCAATCAATCAATCAATCAAT 278
Db 206 AGGAGTTTTCATTTCCAGAGGTCAGTCAACATGAATCTAGTCAATCAATCAATCAATCAAT 265
QY 279 GGGACCGCTGAAAATTTCTTCACTGAAGAGGTGGACTCCCGGAAAAATTCACCAAGGAGG 338
Db 266 GGGACCGCTGAAAATTTCTTCACTGAAGAGGTGGACTCTCGAAAAATTCACCAAGGAGG 325
QY 339 GAAAATCCAGATGAAACTTTTGAGAAATTTGAAGAGCTAGGCGCTTTTTCGGCTGCAAGT 398
Db 326 AAAATCCAGTTCACACCTTAGAGAGTTGAGAGAGCTCGGACTTTTTCGGCATACAGGT 385
QY 399 CCGAGAGATATGGTGGCTGGCTGCTTCTCAACACATCTACTCTAGAGCTAGGAGGAGAT 458
Db 386 CCCAGAGAAATATGGTGGCTGGCTTCTCCAAACAGATCTATGCTCGGCTCGGAGGAGAT 445
QY 459 CATCAGCATGGTGGTCCATCACTGTGACCTTGGAGCCCTGGAGCCACAGCTATTCGCTCAA 518
Db 459 CATCAGCATGGTGGTCCATCACTGTGACCTTGGAGCCCTGGAGCCACAGCTATTCGCTCAA 518

Db	446	CATCAGCCTGGATGCCTCTATACGGTGACCGTGGACCGCTGGCAGCACACCGGCTATCGGCCTCAA	505	
Qy	519	GGGAGCATCTTGGCTGGCACTGAGGAGCAGAAAGCAAAATACTTGCCTAAACTGGCGTC	578	
Db	506	GGGATATCTTGGTCGGCAATGAGGAACAGAAAGGCGCAAGTATCTGCCCAAACTGTCGTC	565	
Qy	579	CGGGAGCACAATTCAGCGCTTCTGCCTCAGGAGCAGCCAGCCAGTGGGAGGATGCGACGCC	638	
Db	566	TGGGAAACATCGCAGCCTTCTGCCTAACAGAGCAGCCAGTGGGAGTGTACTCGGCT	625	
Qy	639	AATCCGAGCAGACCACTAAGTGAAGACAAGAGCAGCTACATCCTCAATGGCTCAA	698	
Db	626	CATCCAGCAGCAGCTACGTTAAGTGAAGATAAGAGTACTTCATCTCAATGGCTCAA	685	
Qy	699	GGTCTGGATCTAATGGAGACTTGGCCAAATATTTTACTGTGTTTGCRAAGACTGAGGT	758	
Db	686	GGTTTGGATCACCATGGAGACTGGCCAAATATTTTACTGTGTTTGCRAAAACCGAGGT	745	
Qy	759	CGTTGATCTCATGATCAGTGAAGAGCAAAATCACAGCATTCATAGTAGAAGAGACTT	818	
Db	746	GGTCGATCTCATGATGTTGGAACACAGACAAATGACCGCATTCATAGTAGAAGAGACTT	805	
Qy	819	TSGTGGAGTCTAATGGGAAACCCGAGATAAATTAGGCATTCGGGGCTCCACACTTG	878	
Db	806	CGCGGAATCACTAATGGGAAACCTGAAGATAAATTAGGCATTCGAGGCTCCACACATG	865	
Qy	879	TGAAGTCCATTTTGAACACCAAGATACCTGTGGAACATCCTTGGAGAGTTCGGAGA	938	
Db	866	TGAGTCCATTTTGAATAATACAGAGTCCCTGTCGGAATGTCTTGGAGAAGTTGGAGG	925	
Qy	939	TGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGTTCAGCATGGGCGCTGCT	998	
Db	926	GGGCTTAAAGTAGCATGAACATCCTGAACAGTGGCGATTCAGCATGGGCGCTGCT	985	
Qy	999	GGCTGGGCTCTCAAGAGATGTGATGAATGACTGCTGAGTACCGCTGCACAAGAAACA	1058	
Db	986	GGCCGGGATGCTGAAGAGAGCTGTCGAAGTGGGTTGATCCAGGAAAGTTGGCTCATGGCTCA	1105	
Qy	1059	GTTTAAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGACATGATGGCTCA	1118	
Db	1046	GTTCAACAGGATCTCAGTGAATTTGGGTTGATCCAGGAAAGTTGGCTCATGGCTCA	1105	
Qy	1119	GAAGGCTTACGTCATGGAGATGATGACCTACTCACAGCGGATGCTGGACCAACCTGG	1178	
Db	1106	GAAGGATACGTAATGGAGATGATGGCTTACCTTACGCGGATGCTAGACCAACAGG	1165	
Qy	1179	CTTTCGGAGTCTCATCGAGGAGCATGGTGAAGTGTTCAGCTCCGAGGCGGCTG	1238	
Db	1166	ATTTCCGAGTCTCTATCGAGGAGCAGCATGTTGAGTGTTCAGCTGAGGCTGCTG	1225	
Qy	1239	CGAGTGTGAGTGGGCGCTGCAGATCCTCGGGGCTTGGGCTACACAAGGAGTATCC	1298	
Db	1226	CGAGTGTGAGGAGGCTCTGCAGATCCTTGGGGCTCAGGCTACATGAAGGACTACCC	1285	
Qy	1299	GTACGAGCGCATCTCAGTGCAGCCGAGCATCTCTCATCTTCGAGGGAACCAATGAGAT	1358	
Db	1286	CTATGAGCGCATCTCAGTGCAGTGCAGGATTTGGGTTGATCCAGGAAAGTTGGCTCATGGCTCA	1345	
Qy	1359	TCTCCGAGTGTACATCGCCCTGACGGGCTGCAGCATGCGCGCCGATCCTCTGACTACGAG	1418	
Db	1346	TCTTCGGTGTTCATTTGCCCTGCAGGCGCTGCAGCATGCTGGACGATCTGACCTGAG	1405	
Qy	1419	GATCCATGAGCTTAACAGCGCCAAAGTAGACAGCATGATGATACCGTTGGCGGAGGCT	1478	
Db	1406	GATCAAGAGCTTAAAGTGGCAATGTGACACAGTCATGAGAGAGTGTGGCGAACT	1465	
Qy	1479	TCGGGACTCCTCGGGCGCAACTGTGGACCTGGGCTGCAGCAACCAACATGGATGTGCA	1538	
Db	1466	TCGGGATCTCTGGGCGCAACTGTGGACCTGGGCTTAACAGCGGATCTTGGTGTGCTCA	1525	
Qy	1539	CCCCAGTCTTGGGAGCAGTGCCCAACAAAGTTTGGAGAGACACCTACTGCTTGGCGGAC	1598	
Db	1526	CCCCAGCCTTGGGAGCAGTGCCCAACAAAGCTTGGAGAAATGTCCATTTTGGCGGAC	1585	
Qy	1599	CGTGAGACACTGCTGCTCCGCTTTGGCAAGACCATCATGGAGGAGCAGCTGGTACTGAA	1658	
Db	1586	TGTTGAGACCTGTTACTCCGCTTTTGGAAAGAACATTTAGAGAAACAGCTGGTGTAAA	1645	
Qy	1659	GGGGTGGCCAAACATCTCATCAACTGTATGGGATGAGCGGCTGCTGCGGGCCAG	1718	
Db	1646	GGGGTAGCCAAAGTCTCATCAACTGTATGGCATGACTGCTGCTGCGAGAGCCAG	1705	
Qy	1719	CGGCTCCATCCGATTTGGGCTCCGCAAGCACGACGAGGTTCTTCTTGGGCAACACACTT	1778	
Db	1706	CGGCTCCATCAAGATTTGGGCTCAGAAATCATGATCAGAGGTTCTTCTTGGGCAACACTT	1765	
Qy	1779	CTGCTGGAGGCTTACTTTCAGAAATCTCTTTCAGCCTCTCTAGCTGAGCAGTATGCTCC	1838	
Db	1766	CTGCTGGAGGCTTATTTTCCAGAAATCTCTTTCAGCCTCTCTAGCCTGTCAGAAAGCGCTCC	1825	
Qy	1839	AGAAACCTTAGATGAGCAGATTAAGAAAGTTCAGCAGAGATCCTTGAAGACGAGCCTA	1898	
Db	1826	AGAAACCTTGGATGAGCAGATTAAGAAAGTGTCTCGGAGATCCTTGAAGACGAGCCTA	1885	
Qy	1899	TATCTGTGCGCCACCTCTGACAGGACATGCTGAGCAGGGGACA	1943	
Db	1886	TATCTGTGCGCCACCTCTGACAGAGCATCTTGGGCTGGAACCA	1930	
RESULT 4	AK028571	2906 bp	linear	HTC 05-DBC-2002
LOCUS	AK028571	Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:473240K02 product:VERY-LONG-CHAIN ACYL-COA		
DEFINITION	AK028571	DEHYDROGENASE VLCAD homolog [Homo sapiens], full insert sequence.		
ACCESSION	AK028571	GI:26080919		
VERSION	AK028571.1			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	1	Carninci, P. and Hayashizaki, Y.		
AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitzunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
AUTHORS	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer			
TITLE	Genome Res. 10 (11), 1757-1771 (2000)			
JOURNAL	20530913			
MEDLINE	11076861			
PUBMED	11076861			
REFERENCE	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Glissi, C., King, B., Kochiwa, H.,		


```

RESULT 6
BM808820
LOCUS
DEFINITION AGENCOURT_6582944 NIH_MGC_41 Homo sapiens linear EST 05-MAR-2002
5' mRNA sequence.
ACCESSION BM808820
VERSION BM808820.1 GI:19125643
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: DCTD/DBP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM1979 row: 1 column: 23
High quality sequence stop: 663.
FEATURES
Location/Qualifiers
1..1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5471398"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/notes="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 287 a 253 c 296 g 232 t 3 others
ORIGIN
Query Match 38.9%; Score 952.8; DB 12; Length 1071;
Best Local Similarity 96.7%; Pred. No. 2.2e-198;
Matches 1035; Conservative 0; Mismatches 29; Indels 6; Gaps 6;
QY 46 GCGTGGGGAACATCGGGCAGCATGACGGCTCGGGCTCTTCTCGGCACACGCGCTGCG 105.
1 1
DB 2 GACGAGGGAACATCGGGCAGCATGACGGCTCGGGCTCTTCTCGGCACACGCGCTGCG 61
QY 106 GCTCGTGCCTCGGGGCTGTGGTGTCTCTACCGGAACCGCGGCTACTGGCACCAGC 165
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 62 GCTCGTGCCTCGGGGCTGTGGTGTCTCTACCGGAACCGCGGCTACTGGCACCAGC 121
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 166 CCGCCTGTACGAGCTTCGCCAAGAGCTTTTCCTAGGCAAAATCAAGAAGAAAGAGTT 225
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 122 CCGCCTGTACGAGCTTCGCCAAGAGCTTTTCCTAGGCAAAATCAAGAAGAAAGAGTT 181
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 226 TTCCCATTTCCAGAAAGTTAGCCAAAGTAACTTAATGAATCAATCAGTTCTTGGGACCC 285
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 182 TTCCCATTTCCAGAAAGTTAGCCAAAGTAACTTAATGAATCAATCAGTTCTTGGGACCC 241
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 286 GTGGAAAATTTCTTCACTGAAGAGGTGGACTCCCGAAAAATTTGACGAGAGGGAATC 345
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 242 GTGGAAAATTTCTTCACTGAAGAGGTGGACTCCCGAAAAATTTGACGAGAGGGAATC 301
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 346 CCAGATGAACCTTTGGAGAAATTTGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAA 405
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

Db 302 CCAGATGAACCTTTGGAGAAATTTGAAGAGCCTAGGCGCTTTTGGGCTGCAAGTCCAGAA 361
QY 406 GAATATGGTGGCCTGGCCTTCTCCACACACATGTAAGAGCTAGGGAGATCATCAGC 465
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 362 GAATATGGTGGCCTGGCCTTCTCCAAACACATGTAAGAGCTAGGGAGATCATCAGC 421
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 466 ATGGATGGGTCCCATCACTGTGACCCCTGGCAGGCACACAGCTATTTGGCCCTCAAGGGATC 525
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 422 ATGGATGGGTCCCATCACTGTGACCCCTGGCAGGCACACAGCTATTTGGCCCTCAAGGGATC 481
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 526 ATCTTGGCTGGCCTAGGAGCAGAAAGCCAAATCTTGCCTAAACTTGGCCTCCGGGGAG 585
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 482 ATCTTGGCTGGCCTAGGAGCAGAAAGCCAAATCTTGCCTAAACTTGGCCTCCGGGGAG 541
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 586 CACATTGCAAGCCTTCTGCTCAGGAGCCAGCAGTGGGAGGATGACAGCTCAATCCGG 645
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 542 CACATTGCAAGCCTTCTGCTCAGGAGCCAGTGGGAGGATGACAGCTCAATCCGG 601
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 646 AGCAGAGCCACACTAAGTGAAGACAAAGCACTACATCCTCAATGGCTCCAAGGTCTGG 705
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 602 AGCAGAGCCACACTAAGTGAAGACAAAGCACTACATCCTCAATGGCTCCAAGGTCTGG 661
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 706 ATTACTAATGGAGGACTGGCCAAATATTTTACTGTGTTTGCANAGACTGAGTCTGAT 765
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 662 ATTACTAATGGAGGACTGGCCAAATATTTTACTGTGTTTGCANAGACTGAGTCTGAT 721
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 766 TCTGATGATCAGTGAAGACAAATACAGCATTCATAGTAGAAGAGACTTTGGTGGGA 825
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 722 TCTGATGATCAGTGAAGACAAATACAGCATTCATAGTAGAAGAGACTTTGGGGA 781
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 826 GTCACTAATGGAAACCCGAAAGATAATTTAGGCAATTC-GGGGCTCCAACACTTGTGAAGT 884
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 782 GTCACTAATGGAAACCCGAAAGATAATTTAGGCAATTC-GGGGCTCCAACACTTGTGAAGT 841
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 885 CATTTTTGAACACCAAGATACCTGTGGAACAT-CCTTGGAGAGTGGAGAT-CGG 942
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 842 CATTTTTGAACACCCAGATACCTGTGGAACATCCCTTGGAGAGTCCGAGATGGG 901
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 943 TTTAAGTGGCCATGAACATCCTCAACAGCGG-CCGGTTTCAGCAT-GGGCAGGCTCGTGG 1000
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 902 TTTAAGTGGCCATGAACATCCTCAACAGCGGCGGTTCANCATGGGCGAGCTCNTGG 961
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 1001 CTGGGCTGCTCAAGAGATTGATTGAATGAGTGTGAGTACGCTGCACAGGAAACAGT 1060
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 962 CTGGGCTGCTCAAGAGATTGATTGAATGAGTGTGAGTACGCTGCACAGGAAACAGT 1021
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 1061 TTACACAGAGCTCAGTGAATTT-GGATTGATTCAGGAGAAATTTGGCACT 1109
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 1022 TTAACAAGAGCTCAATGGATTGGGATTGATTCGGGAGAAATTTTGGCT 1071
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 7
AL539220
LOCUS
DEFINITION AL539220 Homo sapiens FETAL BRAIN Homo sapiens cdna clone
CS0DF034YA11 5-PRIME, mRNA sequence.
ACCESSION AL539220
VERSION AL539220.2 GI:31263790
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
L1,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12868235.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of

```

Invitrogen. This sequence belongs to sequence cluster 2225.f For more information about this cluster, see

http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF034AA06Qp1&cluster=2225.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF034AA06Qp1.

FEATURES

source

Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF034Y11"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
320 a 266 c 306 g 244 t 65 others

BASE COUNT

ORIGIN

Query Match 38.4%; Score 940.4; DB 9; Length 1201;
Best Local Similarity 94.7%; Pred. No. 1.2e-195;
Matches 983; Conservative 27; Mismatches 23; Indels 5; Gaps 4;

QY 32 GAGACTGAGGCTGAGGCTGGGGAACATCGGCGCAGCATGAGCGCTGGCG-GCTCTTCCG 90
DB 63 GGGATTAGGCTGAGGCTGGGGAACATCGGCGCAGCATGAGCGCTGGCGGCTCTTCCG 122
QY 91 CGCACCGCTCGGCTCGCTCGCGGCTGCTGCTCTACCGCGAACCAGCGCG 150
DB 123 CGCACCGCTCGMCTCGCTCGCGGCTGCTGCTCTACCGCGAACCAGCGCG 182
QY 151 CTACTGCCACCGCCGCTGTAGAGCTTTCGCCAAAGAGCTTTCCTAGGCAATC 210
DB 183 CTACTGCCACCGCCGCTGTAGAGCTTTCGCCAAAGAGCTTTCCTAGGCAATC 242
QY 211 RAGAGAAGAGTTCCTCCATTTCCAGAGTTAGCCAAAGTAACTTAATGAATCAAT 270
DB 243 RAGAGAAGAGTTCCTCCATTTCCAGAGTTAGCCAAAGTAACTTAATGAATCAAT 302
QY 271 CAGTTCTTGGACCGCTGGAAAAATCTTCTACTGAAGGTGGACTCCGAAAAATTCAC 330
DB 303 CAGTTCTTGGACCGCTGGAAAAATCTTCTACTGAAGGTGGACTCCGAAAAATTCAC 362
QY 331 CAGGAAGGAATCCAGATGAACTTTGGAGAAATTAAGAGCTAGGCTTTTGGG 390
DB 363 CAGGAAGGAATCCAGATGAACTTTGGAGAAATTAAGAGCTAGGCTTTTGGG 422
QY 391 CTGCAAGTCCCAAGAAATATGTTGGCTGGCTTCTCCACACATGATCTACAGACTA 450
DB 423 CTGCAAGTCCCAAGAAATATGTTGGCTGGCTTCTCCACACATGATCTACAGACTA 482
QY 451 GGGAGATCATCAGATGGTGGTTCATCTGACCTGGAGCGCCACAGGCTATT 510
DB 483 GGGAGATCATCAGATGGTGGTTCATCTGACCTGGAGCGCCACAGGCTATT 542
QY 511 GGCCTCAAGGGATCATCTTGGCTGGCTGAGGAGCAGAAAGCCAAATACTTGCCTAAA 570
DB 543 GGCCTCAAGGGATCATCTTGGCTGGCTGAGGAGCAGAAAGCCAAATACTTGCCTAAA 602
QY 571 CTGGCTCGGGGAGACATTCGAGCTTCTCCCTCAGCGCCAGCCAGTGGAGCGAT 630
DB 603 CTGGCTCGGGGAGACATTCGAGCTTCTCCCTCAGCGCCAGCCAGTGGAGCGAT 662
QY 631 GCAGCTCAATCCGAGCAGACACTAAGTGAAGCAAGAACACTACATCTCAAT 690
DB 663 GCAGCTCAATCCGAGCAGACACTAAGTGAAGCAAGAACACTACATCTCAAT 722
QY 691 GGCTCAAGGTCTGGATTACTAATGGAGGACTGGCCAAATTTTACTCTGTTTCAAG 750

DB 723 GGCTCAAGGTCTGGATTACTTAATGGAGGACTGGCCAAATATTTTACTGTTTGCAGAG 782
QY 751 ACTGAGGTCTGTTGATCTGATGGATCAGTGAACACAAATATCAGCATTCATAGTAA 810
DB 783 ACTGAGGTCTGTTGATCTGATGGATCAGTGAACACAAATATCAGCATTCATAGTAA 842
QY 811 AGAGACTTTGGTGGAGTCACTAATGGGAAACCCCAAGATAAATAGGCATTCGGGGCTCC 870
DB 843 AGAGACTTTGGTGGAGTCACTAATGGGAAACCCCAAGATAAATAGGCATTCGGGGCTCC 902
QY 871 AACACTTGTGAGTCCATTTTCAAAACACCAAGATACCTGTGGAACATCCTTTGGAGAG 930
DB 903 AASASTTGTGAATCSTATTTTGAACACACSAAGATACSTGTGGAACATCCTTTGAGAG 962
QY 931 GTCGGAGATGGTTTAAAGTGGCCATGAACATCTCTCAACAGCGCGGCTTCAGCATGGGC 990
DB 963 GTCGGAGATGGTTTAAAGKGGSCATGAACATCTCTCAACAGCGCGGCTTCAGCATGGGC 1022
QY 991 AGCGTCTGGCTGGCTGCTCAAGAGATGATGAATGAATGCTGCTGAGTACGCTGCACACA 1050
DB 1023 ASC-KCGTGGCTGGCTG-TYAARAGATTATTTGAAATRMWT--YTGAGTAGCCTGCCCA 1078
QY 1051 AGSAAACAGCTTTAACAAG 1068
DB 1079 AGGAAGTTTANAARAG 1096

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 2225.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DE007AF12Qp1&cluster=2225.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DE007AF12Qp1.

Location/Qualifiers

1. .1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DE007YK23"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

BASE COUNT 290 a 284 c 337 g 240 t 50 others

ORIGIN

Query Match

38.2%; Score 937.8; DB 13; Length 1201;

```

Best Local Similarity 97.6%; Pred. No. 4,40-195;
Matches 983; Conservative 9; Mismatches 11; Indels 4; Gaps 4;
QY 778 GTGAAGACAAAATACACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGG 137
Db 66 GTGAAGACAAAATACACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGG 125
QY 838 AAACCCGAGATAAATAGGCATTCGGGGCTCCCAACACTTGTGAAGTCCATTTCAGAA 897
Db 126 AAACCCGAGATAAATAGGCATTCGGGGCTCCCAACACTTGTGAAGTCCATTTCAGAA 185
QY 898 ACCAAGATACCTGTGGAAGACATCCTTGGAGAGTGGAGATGGGTTAAAGTGGCCAT 957
Db 186 ACCAAGATACCTGTGGAAGACATCCTTGGAGAGTGGAGATGGGTTAAAGTGGCCAT 245
QY 958 AACATCCTCAACAGCGCGCGTTCAGCATGGGACGCTGCTGGCTGCTCAAGACA 1017
Db 246 AACATCCTCAACAGCGCGCGTTCAGCATGGGACGCTGCTGGCTGCTCAAGACA 305
QY 1018 TTGATTGAATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAAGAGGCTCAAT 1077
Db 306 TTGATTGAATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAAGAGGCTCAAT 365
QY 1078 GAATTTGGATTCATTACAGAGAAATTTGCACATGATGGCTCAAGAGCTTACGTATG 1137
Db 366 GAATTTGGATTCATTACAGAGAAATTTGCACATGATGGCTCAAGAGCTTACGTATG 425
QY 1138 AGTATGACCTACTCACAGCAGGATGCTGGACCAACCTGGCTTCCCGACTGCTCCATC 1197
Db 426 AGTATGACCTACTCACAGCAGGATGCTGGACCAACCTGGCTTCCCGACTGCTCCATC 485
QY 1198 GAGGACGACCATGGTGAAGGTGTTACAGTCCGAGCGCGCTGGCAGTGTGAGTAGGCG 1257
Db 486 GAGGACGACCATGGTGAAGGTGTTACAGTCCGAGCGCGCTGGCAGTGTGAGTAGGCG 545
QY 1258 GTGACAGTCTCGGGGGCTTGGGCTTACACAGGGACTATCCGTACGACGCGCATACTCGT 1317
Db 546 GTGACAGTCTCGGGGGCTTGGGCTTACACAGGGACTATCCGTACGACGCGCATACTCGT 605
QY 1318 GACACCCGATCTCTCATCTTCGAGGAACCAATGATATCTCCGGATGTACATGCC 1377
Db 606 GACACCCGATCTCTCATCTTCGAGGAACCAATGATATCTCCGGATGTACATGCC 665
QY 1378 CTGACGGTCTGCAATGTC-CGGCGCGCATCTGACTACCAAGGATCCATGAGCTTAAACA 1436
Db 666 CTGACGGTCTGCAATGTC-CGGCGCGCATCTGACTACCAAGGATCCATGAGCTTAAACA 725
QY 1437 GCCCAAGTGAACACATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGG 1496
Db 726 GCCCAAGTGAACACATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGG 784
QY 1497 AACTGTGACCTGGGGCTGACAGCAACCATGGAGTTGTGCACCCAGTCTTGGGACAG 1556
Db 785 AACTGTGACCTGGGGCTGACAGCAACCATGGAGTTGTGCACCCAGTCTTGGGACAG 844
QY 1557 TGCCCAAGTTTGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACACTGCTGCT 1616
Db 845 TGCCCAAGTTTGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACACTGCTGCT 904
QY 1617 CCGCTTTGGCAAGACCATCATGGAGGACGCTGGTACTCAAGCGGGTGGCCAAATCCT 1676
Db 905 CCGCTTTGGCAAGACCATCATGGAGGACGCTGGTACTCAAGCGGGTGGCCAAATCCT 964
QY 1677 CATCAACGTGATGGCATGACGGCGGTGCTGTGCGGGGCGAGCGCTCCATCCGCATTGG 1736
Db 965 CATCAACGTGATGGCATGACGGCGGTGCTGTGCGGGGCGAGCGCTCCATCCGCATTGG 1024
QY 1737 GCTCCGCAACACACACAGGTTCTTCCTGGGCAACACACTTCTGGG 1783
Db 1025 GST-CGCAACMACGA-CACGAGGKTTTYYTGGCAACACTTTTSGT 1069

```

AL567735/c
LOCUS
DEFINITION AL567735 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone
CS0DF034YALL 3-PRIME, mRNA sequence.
ACCESSION AL567735
VERSION AL567735.2 GI:31290607
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12921390.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
2225.f for more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF034AA06NP1&cluster=2225.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DF034AA06NP1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF034YALL"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 271 a 302 c 295 g 270 t 63 others
ORIGIN
Query Match 35.9%; Score 879.2; DB 9; Length 1201;
Best Local Similarity 96.5%; Pred. No. 3.2e-182;
Matches 953; Conservative 17; Mismatches 11; Indels 7; Gaps 7;
QY 1316 GTGACACCCGATCTCTCTATCTTCGAGGGGAACCAATGAGATTCCTCCGATGTACATCG 1375
Db 983 SKGAMACCCGATCTCTCTATCTTCGAGGGGAACCAATGAGATTCCTCCGATGTACATCG 925
QY 1376 CCTGACGGGTCTGC-AGCATCGCGCGCATCTCTGACTACCAAGGATCCATGAGCTT-AA 1433
Db 924 GCTGACGGGTCTGCAAGCATGCCGCGCATCTCTGACTACCAAGGATCCATGAGCTTAAA 865
QY 1434 ACAGGCCAAAGTGAACAGCATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGG 1493
Db 864 ACAGGCCAAAGTGAACAGCATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGG 805
QY 1494 CGGAACCTGGGACCTGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGTCTTGGGGA 1553
Db 804 CGGAACCTGGGACCTGGGCTGACAGGCAACCATGGAGTTGTGCACCCATGCTTGGGGA 745
QY 1554 CAGTCCCAACAGTTTGAAGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACACTGCT 1613
Db 744 CAGTCCCAACAGTTTGAAGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACACTGCT 685
QY 1614 GCTCCGCTTTGGCAAGACCATCATGGAGGACGCTGGTACTGAAGCGGGTGGCCAAACAT 1673
Db 684 -STCCGCTTTGGCAAGACCATCATGGAGGACGCTGGTACTGAAGCGGGTGGCCAAACAT 626

FEATURES
source

QY	1674	CCTCATCAACCTGTATGCGATGACGGCGCGTGTGTCGGGGGACGCGCTCATCCGCAAT	1733
Db	625	CCTCATCAACCTGTATGCGATGACGGCGGT - STGTGCGGGGACGCGCTCATCCGCAAT	567
QY	1734	TGGGCTCCGCAACACGACGAGGTTCTCTTTGGCCAAACCTTCTGCGTGGAGCTTA	1793
Db	566	T - GSGTCCGCAACACGACGAGGTTCTCTTTGGCCAAACCTTCTGCGTGGAGCTTA	508
QY	1794	CTTGCGAAGATCTCTTCAGCCCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTAGATGA	1853
Db	507	CTTGCGAAGATCTCTTCAGCCCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTAGATGA	448
QY	1854	CGAGATTAGAAAGTGTCGCCAGAGATCCTTTGAGAGCCAGCCTATATCTGTGCCACCC	1913
Db	447	GCAGATTAGAAAGTGTCGCCAGAGATCCTTTGAGAGCCAGCCTATATCTGTGCCACCC	388
QY	1914	TCGTGACAGGACATGCTGAGCGAGGGAGAGTGTCCCTGTACCGCCGCCCTACCCA	1973
Db	387	TDGTGACAGGACATGCTGAGCGAGGGAGAGTGTCCCTGTACCGCCGCCCTACCCA	329
QY	1974	TGGCCGCTGCGATGACATGTTACTCTTTTTCAGAGGTTGCGGATATATCAGAGTT	2033
Db	328	TGGCCGCTGCGATGACATGTTACTCTTTTTCAGAGGTTGCGGATATATCAGAGTT	269
QY	2034	AAGCCTTTTGTTCCTCCCTCTGACCTGAAGGTTGTTCGCTGGCCTGGGAGAGCCTCTTC	2093
Db	268	AAGCCTTTTGTTCCTCCCTCTGACCTGAAGGTTGTTCGCTGGCCTGGGAGAGCCTCTTC	209
QY	2094	CAGGTTTTTGACCTGCGAGGAGTCTCTCTTAACAGAGCAATCAGAGCTTCTGAATGAGCC	2153
Db	208	CAGGTTTTTGACCTGCGAGGAGTCTCTCTTAACAGAGCAATCAGAGCTTCTGAATGAGCC	149
QY	2154	GGAGAGAGAGAAATGGAATCTCTGACCCCTGGAACTGGCGGTTATCTGGTCAATTGAGGAG	2213
Db	148	GGAGAGAGAGAAATGGAATCTCTGACCCCTGGAACTGGCGGTTATCTGGTCAATTGAGGAG	89
QY	2214	ACACCATAGTGGAACTGGGGCTTATGCTGTGCTCCAGAGGTGTGAGGTGGTGGGGAC	2273
Db	88	ACMCATAGTGGAACTGGGGCTTATGCTGTGCTCCAGAGGTGTGAGGTGGTGGGGAC	29
QY	2274	CTGTGTCAGGTGTGGATAGCATTCTCTG 2301	
Db	28	KTGTGTCGAGGTGTGATGATKKTCTTTTK 1	
RESULT 10	AL549834	969 bp	linear
LOCUS	AL549834	969 bp	linear
DEFINITION	AL549834 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA		
ACCESSION	AL549834		
VERSION	AL549834.2		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Li.W.B., Gruber,C., Jessee,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12886203.		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	Library was constructed by Life Technologies, a division of		
	Invitrogen. This sequence belongs to sequence cluster 2225.f For		
	more information about this cluster, see		
	http://www.genoscope.cns.fr/		
	cgi-bin/cluster.cgi?seq=CS001054AH090Plscluster=2225.f. Contact :		
	Feng Liang Email : fliang@life.techn.com URL :		
	http://fulllength.invitrogen.com/Invitrogen Corporation 1600		

```

|||||
847 CAGCATTCATAGTAAAGACACTTGTGGTGGAGTCACTAATGGGAACCCCAAGATAAAT 906
|||||
854 TAGGCATTCGGGGTCCCAACACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGG 913
|||||
907 TAGGCATTCGGGGT-CMACACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGG 965
|||||
914 AAAA 917
|||||
966 AAAA 969

RESULT 11
BM552756
LOCUS
DEFINITION
AGENCOURT_6573876 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5477224
5', mRNA sequence.
BM552756
VERSION
BM552756.1 GI:18790885
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1010)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1994 row: o column: 17
High quality sequence stop: 711.
Location/Qualifiers
1. .1010
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5477224"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_98"
/note="Organ: brain; Vector: pOTB7; Site:1: XhoI; Site:2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT
229 a 289 c 285 g 206 t 1 others
ORIGIN

Query Match
Best Local Similarity 35.4%; Score 867.2; DB 12; Length 1010;
Matches 909; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 1042 CCCTGCACAAGCAACAGTTTAAACAAGAGGCTCAGTCAATTTGGATTGATTCAGGAGAAA 1101
|||||
1 GCCTGCACAAGCAACAGTTTAAACAAGAGGCTCAGTCAATTTGGATTGATTCAGGAGAAA 60
|||||
QY 1102 TTTGCATGATGGCTCAGAAGGCTTACCTCATGGAGAGTATGACCTACCTACACAGAGGG 1161
|||||
61 TTTGCATGATGGCTCAGAAGGCTTACCTCATGGAGAGTATGACCTACCTACACAGAGGG 120
|||||
QY 1162 ATGCTGCACCAACCTGCTTCCGAGTCTCCATCGAGCAGCCAGTGGTGAAGGTTC 1221
|||||

```

```

Db 121 ATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGCCATGTTGAAGGTGTTTC 180
QY 1222 AGCTCCGAGAGCCGCTGGCAGTGTGTGAGTGAGGGCTTCAGATCCTCGGGGGCTTGGGC 1281
Db 181 AGCTCCGAGAGCCGCTGGCAGTGTGTGAGTGAGGGCTTCAGATCCTCGGGGGCTTGGGC 240
QY 1282 TACACAAGGAGCTATCCGTACGAGGCGCATACCTGCGTGACACCCGCGATCCTCTCATCTTC 1341
Db 241 TACACAAGGAGCTATCCGTACGAGGCGCATACCTGCGTGACACCCGCGATCCTCTCATCTTC 300
QY 1342 GAGGNAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGTCTCGAGCATCCCGGC 1401
Db 301 GAGGNAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGTCTCGAGCATCCCGGC 360
QY 1402 CGCATCTCTGACTACCAAGATCCAGCTTAAACAGGCGCAAGTGTAGCACAGTCAATGAT 1461
Db 361 CGCATCTCTGACTACCAAGATCCAGCTTAAACAGGCGCAAGTGTAGCACAGTCAATGAT 420
QY 1462 ACCGTTGGCGGAGGCTTCCGGACTTCCCTGGGCGCAACTGTGGACTGGGCTGACAGGC 1521
Db 421 ACCGTTGGCGGAGGCTTCCGGACTTCCCTGGGCGCAACTGTGGACTGGGCTGACAGGC 480
QY 1522 AACCATGAGTGTGCACCCAGCTTTCGGGACAGTGCACAAAGTTTGGAGGAACACACC 1581
Db 481 AACCATGAGTGTGCACCCAGCTTTCGGGACAGTGCACAAAGTTTGGAGGAACACACC 540
QY 1582 TACTGCTTTCGGCGGAGCGGTGGAGACACTGCTGCTCGCTTTGGCAAGACCATCATCGAG 1641
Db 541 TACTGCTTTCGGCGGAGCGGTGGAGACACTGCTGCTCGCTTTGGCAAGACCATCATCGAG 600
QY 1642 GAGCAGCTGGTACTGAAGCGGTGGCAACATCTCATCAACCTGTATGGCATGACGGCC 1701
Db 601 GAGCAGCTGGTACTGAAGCGGTGGCAACATCTCATCAACCTGTATGGCATGACGGCC 660
QY 1702 GTGCTGTGCGGGCGCAGCGCTCCATCGGATTCGGCTCGGCTCCGACACAGCAGCAGGTT 1761
Db 661 GTGCTGTGCGGGCGCAGCGCTCCATCGGATTCGGCTCGGCTCCGACACAGCAGCAGGTT 720
QY 1762 CTCTTGGCAACACCTTCTCGGTGGAGTCTTACTTTCAGAAATCTCTTTCAGCCCTCTCTCAG 1821
Db 721 CTCTTGGCAACACCTTCTCGGTGGAGTCTTACTTTCAGAAATCTCTTTCAGCCCTCTCTCAG 780
QY 1822 CTGGACAAGTATGCTCCAGAAAACCTAGATGAGCAGATTGAAGAGTGTCCCAAGCAGAT - 1880
Db 781 CTGGACAAGTATGCTCCAGAAAACCTAGATGAGCAGATTGAAGAGTGTCCCAAGCAGATC 840
QY 1881 CCTTGAAGAGCGGCTTATATCTGTG-CCACACCTCTTGGACAGGAGATGCTGAGGCGAGG 1939
Db 841 CCTTGAAGAGCGGCTTATATCTGTGCGCCACCTCTTGGACAGGAGATGCTTGAAGCAAG 900
QY 1940 GACAGTGTCCCTCTGCTACCGCCGCGCCCTACCCATGCCCGGCTTGC 1984
Db 901 GGCACATGTCCCTCTGCTACCGCCGCGCCCTTACCCATGCCCGCTTACCCCATGCGC 945

RESULT 12
BM561207
LOCUS
DEFINITION
AGENCOURT_6589858 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5475994
5', mRNA sequence.
BM561207
VERSION
BM561207.1 GI:18806273
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1079)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

```

Tissue Procurement: DCTD/DRP
cDNA Library Preparation by: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
High quality sequence stop: 700.
Location/Qualifiers
1. .1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5475994"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

BASE COUNT 295 a 245 c 311 g 226 t 2 others
ORIGIN
Query Match 35.04; Score 859.4; DB 12; Length 1079;
Best Local Similarity 96.04; Pred. No. 6.9e-178;
Matches 924; Conservative 0; Mismatches 32; Indels 6; Gaps 4;
QY 53 GAACATCGGCAGCATGAGCGGCTCGGGCTCTCTCGGCACACCGGCTCGGCTCGTG 112
DB 1 GAACATCGGCAGCATGAGCGGCTCGGGCTCTCTCTCGGCACACCGGCTCGGCTCGTG 60
QY 113 CCTCGCGGGTCTGGTGGTCTTACCGCGAACCAGCGGGCTTACTGCGGACCGCGGCTG 172
DB 61 CCTCGCGGGTCTGGTGGTCTTACCGCGAACCAGCGGGCTTACTGCGGACCGCGGCTG 120
QY 173 TACGAGCTTTCGCAAGAGAGCTTTCTTAGGCAAAATCAAGAAAGAGTGTTCCTCAT 232
DB 121 TACGAGCTTTCGCAAGAGAGCTTTCTTAGGCAAAATCAAGAAAGAGTGTTCCTCAT 180
QY 233 TTCCAGAGTTAGCAAGATGAATTAATGAATCAATCAATGTTCTGGGACCCCGTGAAA 292
DB 181 TTCCAGAGTTAGCAAGATGAATTAATGAATCAATCAATGTTCTGGGACCCCGTGAAA 240
QY 293 AATTCTTCACTGAAGAGTGGACTCCCGAAATTTGACGAGGAGGAAATCCAGATG 352
DB 241 AATTCTTCACTGAAGAGTGGACTCCCGAAATTTGACGAGGAGGAAATCCAGATG 300
QY 353 AAACCTTTCGAGAAATGAAGAGCTTAGGGCTTTTGGGCTCAAGTCCCAAGAAATATG 412
DB 301 AAACCTTTCGAGAAATGAAGAGCTTAGGGCTTTTGGGCTCAAGTCCCAAGAAATATG 360
QY 413 GTGGCCTGGGCTTCTCCAAACACCATGTACTCAAGACTAGGGGAGATCATCATGATG 472
DB 361 GTGGCCTGGGCTTCTCCAAACACCATGTACTCAAGACTAGGGGAGATCATCATGATG 420
QY 473 GTTCATCACTGTGACCTTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTGG 532
DB 421 GTTCATCACTGTGACCTTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTGG 480
QY 533 CTGGCACTGAGGAGCAGAAAGCAATATCTTGCCTAAATAGTCGCTCCGGGAGCACATG 592
DB 481 CTGGCACTGAGGAGCAGAAAGCAATATCTTGCCTAAATAGTCGCTCCGGGAGCACATG 540
QY 593 CAGCCTTTCGCTCAGGAGCAGCCAGTGGGAGGATGCAGCTCAATCCGAGCAGAG 652
DB 541 CAGCCTTTCGCTCAGGAGCAGCCAGTGGGAGGATGCAGCTCAATCCGAGCAGAG 600

QY 653 CCACACTAAGTGAAGACAAGACACTACATCTCTCAATGGCTCAAGGTCTGGATTACTA 712
DB 601 CCACACTAAGTGAAGACAAGACACTACATCTCTCAATGGCTCAAGGTCTGGATTACTA 660
QY 713 ATGAGGACTGCGCAATATTTTACTCTGTTTGCAAAGACTGAGGTCTGTTGATTCTGATG 772
DB 661 ATGAGGACTGCGCAATATTTTACTCTGTTTGCAAAGACTGAGGTCTGTTGATTCTGATG 720
QY 773 GATCAGTGAAGACAAGACACTACATCTCTCAATGGCTCAAGGTCTGGATTACTA 832
DB 721 GATCAGTGAAGACAAGACACTACATCTCTCAATGGCTCAAGGTCTGGATTACTA 780
QY 833 ATGGAACACCGAAGATTAATTTAGGCATTC-GGGCTCCACACTCTGTGAA-GTCCATTT 890
DB 781 ATGGAACACCGAAGATTAATTTAGGCATTCGCGGGCTCCAAACACTTTGGAAGTCCATTT 840
QY 891 TGAACACACCGAAGATTAATTTAGGCATTCGCGGGCTCCAAACACTTTGGAAGTCCATTT 947
DB 841 TGAACACACCGAAGATTAATTTAGGCATTCGCGGGCTCCAAACACTTTGGAAGTCCATTT 900
QY 948 -GGTGGCCATGAACATCTCTCAACAGCGCGGCTTACGATGGGACGCTCGTGGTGGGC 1006
DB 901 GTGGCCATGGAACATCTCTCAACAGCGCGGCTTAAAGCATGGGACGCGGCTCTGGGC 960
QY 1007 TG 1008
DB 961 TG 962

RESULT 13
AL582366/c
LOCUS
DEFINITION
AL582366 Homo sapiens B CELLS (RAMOS CELL LINE) linear EST 01-JUN-2003
Homo sapiens cDNA clone CS0DL006YK09 3-PRIME, mRNA sequence.
AL582366
VERSION
AL582366.2 GI:31320583
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1206)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12950278.
COMMENT
Contact: Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2225.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL006AF05NP1cluster-2225.f. Contact :
Feng Liang Email : fliang@lifetech.com URL : Invitrogen Corporation 1600
http://fulllength.invitrogen.com/ Parady Avenue Genoscope sequence ID : CS0DL006AF05NP1.
Location/Qualifiers
1. .1206
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL006YK09"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 276 a 306 c 292 g 259 t 73 others

FEATURES
source

ORIGIN

Query Match 34.8%; Score 854; DB 9; Length 1206;
 Best Local Similarity 96.9%; Pred. No. 1.1e-176;
 Matches 901; Conservative 10; Mismatches 15; Indels 4; Gaps 4;

QY 1345 GGAACCAATGATTTCTCCGATGTACATCCCTCGAGGCTCTGACGATCGCGCGC 1404
 DB 930 GGAACAAGARATTTCTCGGARTGTACATYCSCTAMGGGTCTCAASAATACCGT-CGC 872

QY 1405 ATCTGTACTACACGATCCATGACCTTAACAGG-CCAAAGTGAAGACATGATGATAC 1463
 DB 871 ATCTGTACTACCMGAYCCATGAGCTTTAAACAGGCCCAAGTGAAGACATGATGATAC 812

QY 1464 CGTTGGCGGAGGCTTCGGGACTCCCTGGCGCAACTGTGAGCTGGGGTGCAGGCA 1523
 DB 811 CGTTGGCGGAGGCTTCGGRACCTCCCTGGCGCAACTGTGAGCTGGGGTGCAGGCA 752

QY 1524 CCATGGAGTTGTGCACCCAGTCTTGGGACAGTGCACCAAGTTTGGAGAACACCTA 1583
 DB 751 CCATGGAGTTGTGCACCCAGTCTTGGGACAGTGCACCAAGTTTGGAGAACACCTA 692

QY 1584 CTGCTTGGCGGAGGCTGTGAGACACTGTGCTCGCTTGGGACAGTGCACCAAGTTTGGAGAACACCTA 1643
 DB 691 CTGCTTGGCGGAGGCTGTGAGACACTGTGCTCGCTTGGGACAGTGCACCAAGTTTGGAGAACACCTA 632

QY 1644 GCAGCTGGTACTGAAGGGGTGGCCAAACATCTCTCATCAACCTGTATGGCATGACGCGCT 1703
 DB 631 GCAGCTGGTACTGAAGGGGTGGCCAAACATCTCTCATCAACCTGTATGGCATGACGCGCT 572

QY 1704 GCTGTGCGGGCGGAGGCTTCCATCGGCACTGGGCTCCGCAACCGACGAGGTTCT 1763
 DB 571 G-TGTCGGGGCGGAGGCTTCCATCGGCACTGGGCTCCGCAACCGACGAGGTTCT 514

QY 1764 CTTGGCCACACCTTCTGCTGGAAGCTTACTTGGAGATCTCTTCCAGTCTCTCAGCT 1823
 DB 513 CTTGGCCACACCTTCTGCTGGAAGCTTACTTGGAGATCTCTTCCAGTCTCTCAGCT 454

QY 1824 GGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAGTGTCCAGCAGATCCT 1883
 DB 453 GGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAGTGTCCAGCAGATCCT 394

QY 1884 TGAGAGGAGGCTTATCTGTGCCACCCCTCTGACAGGACATGCTGAGGAGGAGCA 1943
 DB 393 TGAGAGGAGGCTTATCTGTGCCACCCCTCTGACAGGACATGCTGAGGAGGAGCA 334

QY 1944 GTGTCCCTGTACCGCGGCGGCTTACCATGGCCGCTTGTGGATGACTTTACTCTTT 2003
 DB 333 GTGTCCCTGTACCGCGGCGGCTTACCATGGCCGCTTGTGGATGACTTTACTCTTT 274

QY 2004 TTTCAGAGGTTTGGGATTATCAGAGTTAAGCTTTTGTTCCTGCTGACCTGAAG 2063
 DB 273 TTTCAGAGGTTTGGGATTATCAGAGTTAAGCTTTTGTTCCTGCTGACCTGAAG 214

QY 2064 GTTGTGCGGCTGGGAGGCTTTCAGGTTTGGAGTGTGCTGAGGAGTGTCTCTA 2123
 DB 213 GTTGTGCGGCTGGGAGGCTTTCAGGTTTGGAGTGTGCTGAGGAGTGTCTCTA 154

QY 2124 ACAGACCATCACAGCTTCTCAACTGAGCGGAGAGAGATGAATGCTGACCCCTG 2183
 DB 153 ACAGACCATCACAGCTTCTCAACTGAGCGGAGAGAGATGAATGCTGACCCCTG 94

QY 2184 GAACCTGGGGGTATTTGGTCAATTGAGGAGACACCATAGTGGAAACTGGGGCTTATGCTG 2243
 DB 93 GAACCTGGGGGTATTTGGTCAATTGAGGAGACACCATAGTGGAAACTGGGGCTTATGCTG 34

QY 2244 CTGCTCCAGGCTGTGAGGTGGTGGGAC 2273
 DB 33 CTGCTCCAGGCTGTGAGGTGGTGGGAC 4

RESULT 14
 BQ938039

LOCUS BQ938039 890 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGENCOURT_894875 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6469922
 5' mRNA sequence.
 ACCESSION BQ938039
 VERSION BQ938039.1 GI:22353517
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 890)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 TITLE NIH-MGC http://imgc.ncl.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLNL1999 row: n column: 03
 High quality sequence stop: 611.
 Location/Qualifiers
 1. .890
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6469922"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo df.
 Average insert size 2.1 kb."
 BASE COUNT 242 a 214 c 248 g 186 t

ORIGIN
 Query Match 34.8%; Score 852.4; DB 13; Length 890;
 Best Local Similarity 99.6%; Pred. No. 2.3e-176;
 Matches 886; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 13 CTGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGCATGAGC 72
 DB 1 CTGCGCGCTAAGAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGCATGAGC 60

QY 73 GGCTGGGGCTCTTCTTGGCGACACCGCTGCGGCTCGCTGCGGGGCTGTGGTGGTC 132
 DB 61 GGCTGGGGCTCTTCTTGGCGACACCGCTGCGGCTCGCTGCGGGGCTGTGGTGGTC 120

QY 133 TCTACCGCGACCGGGGCTACTGCGCACCGCCGCTGTACGAGCTTTCGCCAAGAG 192
 DB 121 TCTACCGCGACCGGGGCTACTGCGCACCGCCGCTGTACGAGCTTTCGCCAAGAG 180

QY 193 CTTTCTTAGGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGCCAAAGAT 252
 DB 181 CTTTCTTAGGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGCCAAAGAT 240

QY 253 GAACCTTAATGAATCAATCAGTTCTTGGACCCGCTGGAAATTTCTTCACTGAAGAGTG 312
 DB 241 GAACCTTAATGAATCAATCAGTTCTTGGACCCGCTGGAAATTTCTTCACTGAAGAGTG 300

QY 313 GACTCCGCAAAATTTGACCAAGGAGGAAATCCAGATGAACCTTTGGAGAAATTTGAAG 372
 DB 301 GACTCCGCAAAATTTGACCAAGGAGGAAATCCAGATGAACCTTTGGAGAAATTTGAAG 360

QY 373 AGCCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGCTGGGCTGTCCCAAC 432
 DB 361 AGCCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGCTGGGCTGTCCCAAC 420

QY 433 ACCATGTACTCAAGACTAGGGGAGATCATCAGATGGATGGTCCATCATCTGTGACCCCTG 492

BASE COUNT		329 a	261 c	309 g	245 t	57 others
ORIGIN						
Query Match		34.8%	Score	852.4	DB 9	Length 1201;
Best Local		Similarity	96.1%	Pred. No.	2.4e-176	
Matches		903	Conservative	12	Mismatches	18
					Indels	7
					Gaps	4
QY	11	TCCCTGGCGCTAAGAGGGAGACTGAGCTGAGCTGGGAAACATCGGAGCATGA	70			
DB	60	TCCCTGGCGCTAAGAGGGAGAACTGAGCTGAGCTGGGAAACATCGGAGCATGA	119			
QY	71	CGGCTGCGG-GCTCTTCTCGCCACACAGGCTCGCTCGCTCGCGGGTCTGGTG	129			
DB	120	CGGCTGCGGNCCTCTCTCGCCACACAGGCTCGCTCGCTCGCGGGTCTGGTG	179			
QY	130	GTCCTACCGGAACCGCGCTACTGCGCACACAGCGCTGTACGAGCTTTCGCCAAA	189			
DB	180	GTCCTACCGGAACCGCGCTACTGCGCACACAGCGCTGTACGAGCTTTCGCCAAA	239			
QY	190	GAGCTTTTCTAGGCAAAATCAAGAAGAAAGATTTTCCCATTTCCAGAAATGAGCAA	249			
DB	240	AGCTTTTCTAGGCAAAATCAAGAAGAAAGATTTTCCCATTTCCAGAAATGAGCAA	299			
QY	250	GATGAACCTTAATGAATCAATCAGTTCTTGGACCCGCGGAAATTTCTTCACTGAAGAG	309			
DB	300	GATGAACCTTAATGAATCAATCAGTTCTTGGACCCGCGGAAATTTCTTCACTGAAGAG	359			
QY	310	GTGGACTCCGAAAAATTTGACGAGAGGAGAAATTTCCAGATGAAACTTTTGAGAAATTTG	369			
DB	360	GTGGACTCCGAAAAATTTGACGAGAGGAGAAATTTCCAGATGAAACTTTTGAGAAATTTG	419			
QY	370	AAGAGCTTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGGCTTCCTCC	429			
DB	420	AAGAGCTTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGGCTTCCTCC	479			
QY	430	AACACCATCTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCATCTGTGACC	489			
DB	480	AACACCATCTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCATCTGTGACC	539			
QY	490	CTGGCAGCCGACAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCTGAGGAGGAG	549			
DB	540	CTGGCAGCCGACAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCTGAGGAGGAG	599			
QY	550	AAGCCCAATATCTTGCTTAACTGGCTCCGGGAGACATTTGCAGCCTTCTGCCTCAG	609			
DB	600	AAGCCCAATATCTTGCTTAACTGGCTCCGGGAGACATTTGCAGCCTTCTGCCTCAG	659			
QY	610	GAGCCAGCCAGTGGGAGCGATGAGCTCAATCCGAGCAGACCATTAAGTGAAGAC	669			
DB	660	GAGCCAGCCAGTGGGAGCGATGAGCTCAATCCGAGCAGACCATTAAGTGAAGAC	719			
QY	670	AAGAACCATCTACATCTCAATGCTCCAGGCTCGATTACTTAATGGAGCTGGCCAAAT	729			
DB	720	AAGAACCATCTACATCTCAATGCTCCAGGCTCGATTACTTAATGGAGCTGGCCAAAT	779			
QY	730	ATTTTACTGTGTTTCAAGACTAGGCTGTTGATCTGATGATGATGATGATGATGATGAT	789			
DB	780	ATTTTACTGTGTTTCAAGACTAGGCTGTTGATCTGATGATGATGATGATGATGATGAT	839			
QY	790	ATCACACATCTACATGAGAAAGACTTTTGGTGGAGTCACTAATGGGAAACCCGAAGAT	849			
DB	840	ATCACACATCTACATGAGAAAGACTTTTGGTGGAGTCACTAATGGGAAACCCGAAGAT	899			
QY	850	AAATTAGGCATTCGGGCTCCAAACACTTTGTAAGTCCATTTTGAACACACAGATACCT	909			
DB	900	AAATTAGGCATTCGGGCTCCAAACACTTTGTAAGTCCATTTTGAACACACAGATACCT	955			
QY	910	GTGGAACATCTTGGAGAGGTCGGAGATGGGTTTAAGG	949			
DB	956	GTGGAACATCTTGGAGAGGTCGGAGATGGGTTTAAGG	993			

DB	421	ACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCAGCTGTGACCCTG	480
QY	493	GCAGCCACAGGCTATTGGCTCAAGGGATCATCTTGGCTGGGACTGAGGAGCAGAAA	552
DB	481	GCAGCCACAGGCTATTGGCTCAAGGGATCATCTTGGCTGGGACTGAGGAGCAGAAA	540
QY	553	GCCAAATACTTGCCTAAACTGGCTCCGGGGAGCAGATTTGAGCCCTTCTGCTCAGGAG	612
DB	541	GCCAAATACTTGCCTAAACTGGCTCCGGGGAGCAGATTTGAGCCCTTCTGCTCAGGAG	600
QY	613	CCAG-CCAGTGGAGGATGAGCCTCAATCCGGAGCAGACACACTAAGTGAAGACAA	671
DB	601	CCAGCCAGTGGAGGATGAGCCTCAATCCGGAGCAGACACACTAAGTGAAGACAA	660
QY	672	GAAGCACTACATCTCAATGGGCTCCAGGCTCGGATTTACTAATGAGGAGTGGCCAAAT	731
DB	661	GAAGCACTACATCTCAATGGGCTCCAGGCTCGGATTTACTAATGAGGAGTGGCCAAAT	720
QY	732	TTTTACTGTTTGCCTAAAGACTGAGTCTGTTGATTTCTGATGATGATGATGATGATGAT	791
DB	721	TTTTACTGTTTGCCTAAAGACTGAGTCTGTTGATTTCTGATGATGATGATGATGATGAT	780
QY	792	CACAGCATTCATAGTGAAGAGACTTT-GGTGGAGTCACTAATGGAAGACCCGAGATA	850
DB	781	CACAGCATTCATAGTGAAGAGACTTTGGTGGAGTCACTAATGGAAGACCCGAGATA	840
QY	851	AATTAGGCACTCGGGCT-CCAACTTTGAGTCACTTTGAAACAC	899
DB	841	AATTAGGCACTCGGGCTCCAACTTTGAGTCACTTTTGAACAC	890
RESULT 15			
AL560956			
LOCUS			
DEFINITION			
AL560956 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED			
Homo sapiens cDNA clone CS0DL006YK09 5-PRIME, mRNA sequence.			
ACCESSION			
AL560956			
VERSION			
AL560956.2 GI:31285085			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1. (bases 1 to 1201)			
Li W.B., Gruber C., Jessee J., and Polayes D.			
Full-length cDNA libraries and normalization			
Unpublished			
On Feb 15, 2001 this sequence version replaced gi:12907913.			
JOURNAL			
COMMENT			
Contact: Genoscope			
Genoscope - Centre National de Sequenage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 2225.f For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0DL006AF05QPL&cluster=2225.f. Contact :			
Feng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
Faraday Avenue Genoscope sequence ID : CS0DL006AF05QPL.			
FEATURES			
source			
1..1201			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0DL006YK09"			
/cell_type="B CELLS (RAMOS CELL LINE)"			
/cell_line="RAMOS CELL LINE"			
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT			
25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo(dT)			
primer. Five prime end enriched, double-strand cDNA was			

Tue Sep 9 10:14:37 2003

us-09-945-326-1.rst

Page 17

Search completed: September 7, 2003, 01:51:41
Job time : 5168.14 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:30:07 ; Search time 62 Seconds
(without alignments)
1375.251 Million cell updates/sec

Title: US-09-945-326-2
Perfect score: 3153
Sequence: 1 MSGCGFLFRTTAAARACRGL.....SQILEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513375 seqs, 137303645 residues

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications.AA.*
1: /cgn2.6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2.6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2.6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2.6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2.6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2.6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2.6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2.6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2.6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2.6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2.6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2.6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2.6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2.6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2.6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2.6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2.6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2.6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3153	100.0	621	10	US-09-945-326-2
2	3153	100.0	621	15	US-10-168-274-24
3	659.5	20.9	646	15	US-10-156-761-10104
4	650	20.6	384	15	US-10-272-419-6
5	615.5	19.5	385	15	US-10-156-761-12560
6	604	19.2	432	12	US-10-354-358-28
7	574.5	18.2	390	15	US-10-156-761-14447
8	562.5	17.8	386	15	US-10-156-761-12809
9	556	17.6	383	15	US-10-156-761-14127
10	530.5	16.8	401	15	US-10-156-761-9454
11	495	15.7	385	15	US-10-156-761-12814
12	492.5	15.6	409	15	US-10-156-761-14657
13	480	15.2	513	12	US-10-181-319-33
14	479	15.2	415	15	US-10-205-823-4
15	452.5	14.4	376	12	US-10-238-075-951
					Sequence 2, Appli
					Sequence 24, Appl
					Sequence 10104, A
					Sequence 6, Appli
					Sequence 12560, A
					Sequence 28, Appl
					Sequence 14447, A
					Sequence 12809, A
					Sequence 14127, A
					Sequence 9454, Ap
					Sequence 12814, A
					Sequence 14657, A
					Sequence 33, Appli
					Sequence 4, Appli
					Sequence 951, App

16	395	12.5	346	10	US-09-925-300-1645	Sequence 1645, Ap
17	382.5	12.1	403	15	US-10-156-761-8886	Sequence 8886, Ap
18	374	11.9	375	15	US-10-156-761-9843	Sequence 9843, Ap
19	367.5	11.7	391	15	US-10-156-761-14069	Sequence 14069, A
20	340	10.8	797	11	US-09-999-314A-2	Sequence 2, Appli
21	339	10.8	413	15	US-10-156-761-8259	Sequence 8259, Ap
22	330	10.5	407	15	US-10-156-761-14142	Sequence 14142, A
23	310	9.8	421	15	US-10-156-761-8920	Sequence 8920, Ap
24	298.5	9.5	601	9	US-09-815-242-5070	Sequence 5070, Ap
25	297.5	9.4	386	15	US-10-156-761-11399	Sequence 11399, A
26	295	9.4	156	11	US-09-999-314A-7	Sequence 7, Appli
27	289	9.2	608	15	US-10-156-761-11924	Sequence 11924, A
28	283	9.0	398	10	US-09-738-626-3821	Sequence 3821, Ap
29	275	8.7	392	15	US-10-156-761-8806	Sequence 8806, Ap
30	260	8.2	393	15	US-10-156-761-14669	Sequence 14669, A
31	254.5	8.1	585	10	US-09-976-059-26	Sequence 26, Appl
32	227.5	7.2	400	12	US-09-769-724-50	Sequence 50, Appl
33	222	7.0	379	15	US-10-156-761-13723	Sequence 13723, A
34	207	6.6	384	15	US-10-156-761-11373	Sequence 11373, A
35	207	6.6	429	15	US-10-156-761-11746	Sequence 11746, A
36	204.5	6.5	376	15	US-10-156-761-14670	Sequence 14670, A
37	203	6.4	1261	15	US-10-156-761-11188	Sequence 11188, A
38	202	6.4	71	9	US-09-864-761-47228	Sequence 47228, A
39	199.5	6.3	549	15	US-10-156-761-9661	Sequence 9661, Ap
40	196.5	6.2	383	15	US-10-156-761-11368	Sequence 11368, A
41	190	6.0	413	15	US-10-156-761-11355	Sequence 11355, A
42	181.5	5.8	553	10	US-09-976-059-25	Sequence 25, Appl
43	178.5	5.7	414	15	US-10-119-651-6	Sequence 6, Appli
44	173	5.5	609	15	US-10-156-761-11187	Sequence 11187, A
45	169.5	5.4	394	15	US-10-156-761-8745	Sequence 8745, Ap

ALIGNMENTS

RESULT 1

US-09-945-326-2
; Sequence 2, Application US/09945326
; Patent No. US20020127880A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Hunter, John Joseph
; TITLE OF INVENTION: 62112, A NOVEL HUMAN DEHYDROGENASE AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MNI-187
; CURRENT APPLICATION NUMBER: US/09/945, 326
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/229,831
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-326-2

Query Match	100.0%	Score 3153	DB 10	Length 621
Best Local Similarity	100.0%	Pred. No. 3.4e-298		
Matches 621	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MSGCGFLFRTTAAARACRGLVVTANRRLLRTPSPVAFAKELFLGKIKKKEVPPFPVS	60	
DB	1	MSGCGFLFRTTAAARACRGLVVTANRRLLRTPSPVAFAKELFLGKIKKKEVPPFPVS	60	
QY	61	QDELNEINQFGLPVEKFFTEEVDSRKIDQEKIPDETLEKLSGLFLGLQVPEEYGLGF	120	
DB	61	QDELNEINQFGLPVEKFFTEEVDSRKIDQEKIPDETLEKLSGLFLGLQVPEEYGLGF	120	
QY	121	SNMTYSRLGEIISDMSGITVTLLAAHQALGLKGIILAGTEEQAKYLPKLASEHTAFCFL	180	
DB	121	SNMTYSRLGEIISDMSGITVTLLAAHQALGLKGIILAGTEEQAKYLPKLASEHTAFCFL	180	

Qy	181	T	E	P	A	S	G	D	A	A	I	R	S	R	A	L	S	E	D	K	H	Y	I	L	N	G	S	K	V	I	T	F	A	K	T	E	V	V	D	S	G	S	V	K	D	240														
Db	181	T	E	P	A	S	G	D	A	A	I	R	S	R	A	L	S	E	D	K	H	Y	I	L	N	G	S	K	V	I	T	F	A	K	T	E	V	V	D	S	G	S	V	K	D	240														
Qy	241	K	I	T	A	F	I	V	E	R	D	E	R	G	G	Y	T	N	G	K	P	E	D	K	I	G	R	S	N	T	C	E	V	H	F	E	N	T	K	I	P	V	E	N	I	L	G	E	V	G	D	F	K	V	A	M	N	I	L	300
Db	241	K	I	T	A	F	I	V	E	R	D	E	R	G	G	Y	T	N	G	K	P	E	D	K	I	G	R	S	N	T	C	E	V	H	F	E	N	T	K	I	P	V	E	N	I	L	G	E	V	G	D	F	K	V	A	M	N	I	L	300
Qy	301	N	S	G	R	S	M	G	S	V	A	G	L	L	K	R	I	E	M	T	A	B	A	C	T	R	K	O	F	N	K	R	L	S	E	F	G	L	I	O	E	K	F	A	L	M	A	O	K	A	Y	V	M	E	S	M	T	360		
Db	301	N	S	G	R	S	M	G	S	V	A	G	L	L	K	R	I	E	M	T	A	B	A	C	T	R	K	O	F	N	K	R	L	S	E	F	G	L	I	O	E	K	F	A	L	M	A	O	K	A	Y	V	M	E	S	M	T	360		
Qy	361	Y	L	T	A	G	M	L	D	Q	P	F	P	D	C	S	I	E	A	M	K	V	F	S	S	E	A	A	Q	C	V	S	E	A	L	Q	I	L	G	I	G	I	T	R	D	P	I	P	E	R	I	L	R	O	T	R	420			
Db	361	Y	L	T	A	G	M	L	D	Q	P	F	P	D	C	S																																												

```

RESULT 2
US-10-168-274-24
; Sequence 24, Application US/10168274
; Publication No. US20030124106A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: HILLMAN, Jennifer
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: LU, Dyung Aina M.
; TITLE OF INVENTION: HUMAN OXIDOREDUCTASE PROTEINS
; FILE REFERENCE: PF-0754 PCT
; CURRENT APPLICATION NUMBER: US/10/168,274
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/172,367
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 621
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030124106A1 5540437CD1
; US-10-168-274-24

```

	Query Match	Best Local Similarity	Matches 621;	Conservative	100.0%;	Score 3153;	DB 15;	Length 621;	Gaps 0;
		100.0%;	0;	Mismatches =8;	Pred. NO.	3.4e-298;			
QY	1	MSGGGLFURTAAARACRGLVVTANRRLLTSPVRAFAKELFLGKIKKEVEFPFVS	60						
DB	1	MSGGGLFURTAAARACRGLVVTANRRLLTSPVRAFAKELFLGKIKKEVEFPFVS	60						
OY	61	ODELNEINQFGPVEKFPTTEVEDSRKDIOEGKIIDEFTLEKLKSGLFCLOVPWEVGGLGF	120						

```

61 QDELNEINFLGPVEKFFTEEVDSRSKIDQEKIPDTELEKLSGLGLQVPEEYIGLGF 120
121 SNTMYSRLLGEIISMDGSIPTVTLAAHQAILGLKGIILAGTEEQAKYVLPKLASGEHIAAFCL 180
121 SNTMYSRLLGEIISMDGSIPTVTLAAHQAILGLKGIILAGTEEQAKYVLPKLASGEHIAAFCL 180
121 SNTMYSRLLGEIISMDGSIPTVTLAAHQAILGLKGIILAGTEEQAKYVLPKLASGEHIAAFCL 180
181 TEPASGSDAASRSTRATLSEDKKKHYLLNGSKWVIITNGGLANIFTVFATKTEVVDSDGSVKD 240
181 TEPASGSDAASRSTRATLSEDKKKHYLLNGSKWVIITNGGLANIFTVFATKTEVVDSDGSVKD 240
241 KITATFIVERDFGGVTNGKPEDKLGRSGNTECVHEFNTKIPVENILGEVGDGFKVAMNII 300
241 KITATFIVERDFGGVTNGKPEDKLGRSGNTECVHEFNTKIPVENILGEVGDGFKVAMNII 300
301 NSGRFSGMSGVAGLLKRLIEMTAEEYACTKQFNKRLSEFGLIOEKFALMAQKAYVMESMT 360
301 NSGRFSGMSGVAGLLKRLIEMTAEEYACTKQFNKRLSEFGLIOEKFALMAQKAYVMESMT 360
361 YLTAGMLDQPFDPDCSTEAAMVKVFSSEAAWCVSEALQIILGGLGYTRDYPYERILLRDR 420
361 YLTAGMLDQPFDPDCSTEAAMVKVFSSEAAWCVSEALQIILGGLGYTRDYPYERILLRDR 420
421 ILLIFEGTNEILRWYIALTGLQHAGRLITRIHELKQAKVSTVMDTVGRRLRDSLGRTVD 480
421 ILLIFEGTNEILRWYIALTGLQHAGRLITRIHELKQAKVSTVMDTVGRRLRDSLGRTVD 480
481 LGLTNGNVGVHPSLADSSANKFEENTCYCGRTVETLLLRFGKTIEMEQVLKRVANILNII 540
481 LGLTNGNVGVHPSLADSSANKFEENTCYCGRTVETLLLRFGKTIEMEQVLKRVANILNII 540
541 YGMTAVILSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSOLDKYPENLDEQIKK 600
541 YGMTAVILSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSOLDKYPENLDEQIKK 600
601 VSQQTLEKRAYICAHPLDRTC 621
601 VSQQTLEKRAYICAHPLDRTC 621

RESULT 3
US-10-156-761-10104
; Sequence 10104, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10104
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10104

Query Match 20.9%; Score 659.5; DB 15; Length 646;
Best Local Similarity 30.6%; Pred. No. 4.3e-55;
Matches 186; Conservative 107; Mismatches 242; Indels 73; Gaps 16;

Qy 38 AFAKELFLGKIKKKEVFPEVVSQDELNEINQFLGPVEKFFTEEVDSRSKIDQEKIPDET 97
Db 34 SFAKELFLGRDLTLHPHMPADEAQORGEFLAKLDFCETEDIDSARIEREARIPEDET 93

```



```

Db 266 YAKGVQERKQFGKAIADFQIOFQMLADMAKMEAAARALTYAAAKSER-GDKDLTFQGA 324
QY 301 MVKVFSEAAQCVCSEALQILGGIGYRDYDYPYERILRDLTRILLIFEGTNEILRMVIA 437
Db 325 AAKCFASDVAMEVTTDAVOLLLGGYGYTRDIPVVERMDAKITQIYEGTNQVQRIVMA 381

RESULT 6
US-10-354-358-28
; Sequence 28, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 32320, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 56428 MOLECULES
; FILE REFERENCE: MPI02-020PIRNMNIM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-354-358-28

Query Match 19.2%; Score 604; DB 12; Length 432;
Best Local Similarity 35.9%; Pred. No. 5.7e-50;
Matches 157; Conservative 81; Mismatches 155; Indels 44; Gaps 14;
QY 19 GLVSTANRRLLRTPPVAFKELFLGKIKKVEFP-FPVSODE--LNEINQ--FLGP 73
Db 3 GLAV-----RLRGSELLR----RNFLTCLSSWKIPPHVSKSSQSEALLNITNNGIHFP 53
QY 74 VEKFTTEVDSRK-----IDEGKIPDETLEKLSLGLFLGLQVPEYGG 117
Db 54 LOTFTDEEMNIKSSVKKFAEQIAPLVSTMDENSKMEKSVIOGLFQOGLMGIEVDPEYGG 113

```

```

QY 118 LG--ESNTMYSRGELIISMDGSIQVTLAAHQAIUGLKGILLAGTBEQKAKYLPKLASGEHI 175
Db 114 TGASFLSTVLV-IEELAKVDASVAVFCBIQNTLINTLRKKGHTBEQKATYLPQLTT-EKV 171
QY 176 AAFCLTEPASGSDAASIRSRTLSEDKKHYYILNGSKVWITNGGLANIFTFVFAKTEVVDSD 235
Db 172 GSFCLSEAGAGSDSFALKTRADKEGD--YYVLNGSKMWISSNAEHAGLFLVMANV-----D 224
QY 236 GSVKDK-ITAFIVERDFGGVTNGKPEDKLGIRGNSNTCEVHFENTKIPVENILGEVGDGFK 294
Db 225 PTIGYKGITSLVDROTGLGHGKPKENKLGRLRASSTCLTFENVKVPEANILGQIGHGYK 284
QY 295 VAMNILNSGRFSMGSVVAGLLKRLLEMTAEYACTRKQENKRLSEFGLIOEKFALMAQAKY 354
Db 285 YATGSLNEGRIGITAAQMLGLAQGCCFDYITPYIKERIQFGKRLDFDQGLQHOHVAVATQLE 344
QY 355 VMESMTYLTAGMLDQGFDPDCSTEAAMVKVFSSEAAWQCVSEALQILGLGLGTRDYPYER 414
Db 345 AARLLTYNARLL-EAGKPFIR-EASMAKYVASEIAGQTTSCIEWMGGVGYTKDYPVEK 402
QY 415 ILRDRILLIFEGTNEI 431
Db 403 YFRDAKIGTIVEGASNI 419

RESULT 7
US-10-156-761-14447
; Sequence 14447, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14447
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14447

Query Match 18.2%; Score 574.5; DB 15; Length 390;
Best Local Similarity 34.4%; Pred. No. 3.6e-47;
Matches 135; Conservative 80; Mismatches 152; Indels 25; Gaps 8;
QY 55 PPEVVSQDELNEINQFLGPVEKFFTEEV--DSRKIDQEGKIPDETLEKLSLGLFLGLOVP 112
Db 6 POPVDRLPTDEARDLISLRVIAQREIAPEAAEGEDAGHPREVFGLLSGLGLPYD 65
QY 113 EEVGGIGFNTMYSRGELIISMDGSITV-----TLAAHQAIUGLKGILLAGTBEQKAKY 165
Db 66 SEYGGDQPYEVYLOALELAA-ARLTGLGVSVHTLACH-----ALAQYGTKEQOQVEH 118
QY 166 LPKLASGEHIAFCLTEPASGSDAASIRSRTLSEDKKHYYILNGSKVWITNGGLANIFTV 225
Db 119 LPAMLGGGLGAYCLSEPSGSDAASLRTRKAV--RDGDDWITIGTKAWITHGDIADFTV 176
QY 226 FAKTEVVDSDGSKVKDKITAFIVERDPGGVTNGKPEDKLGIRGNSNTCEVHFENTKIPVENI 285
Db 177 MAT-----GGEGARGITAFVPGDAEGLHAALPEKKMGKMGKSPTAQVHFQVGRVSDERR 231
QY 286 LGEVGDGFKVAMNILNSGRFSMGSVVAGLLKRLLEMTAEYACTRKQENKRLSEFGLIOEK 345

```

```

Db      232  IGDEGQFAIALSALDSGRGLAACAVGVAQAALDEAVATATGQPGRIADPQGLRFEM 291
Qy      346  FALMAKAYVNESYTLTACHLQDPGPPDCSIRIAMVKVFSSEAAQCSEALQILGGLG 405
Db      292  LADMATQIAGRAL-YLAAARLRDAGR-ESKQAAMAKLLCTDAAMKVTITDAVQVLGGYG 349
Qy      406  YTRDPYPERILRTRILLIFEGTNEILRMVIA 437
Db      350  YTADFPVRYMREAKVLIQVEGTNQIQRMVIA 381

RESULT 8
US-10-156-761-12809
; Sequence 12809, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12809
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12809
Query Match
Best Local Similarity 17.8%; Score 562.5; DB 15; Length 386;
Matches 142; Conservative 72; Mismatches 144; Indels 31; Gaps 10;
Qy      63  ELNEINQFLGPVEKFFTE-EVDSRKID---QEGKIPDETLEKLSGLGLGLOVPEYGG 119
Db      8  ELEELRR---TVEE-FADVVPAPKIGDFYERHETPEYELVREMGRLGFLPFPPEYGG 63
Qy      120  FSNTMYSRLG-----EIIISMDGSIYVTLAAHQAIKGLGIIAGTEEQAKYLPKLAGSEHI 175
Db      64  GD---YLALGIALLEELARVDSVAITLEAGVSLGAMPILHFGTDQAQAEWLPRLCSG 120
Qy      176  AAFCLTEPASGSDAASIRATLSDEKHKHYILNGSKVWITNGG--LANIFTVFAKTEVD 233
Db      121  GAFGLTEPDGSDAGATRTTARLDESNEWVINGTKCFITNSGDTITGLTV---TAVTG 177
Qy      234  SDGSKKITAFIVERDFGVTNGKPEDKLGIRGNTCEVFHENTKIPVENILGEVGDGF 293
Db      178  RKPDKPLISSIIVPSTGPTGPTVAAPYSKVGWNASDTRLSFADVRVPAANLGEQGRG 237
Qy      294  KVAMNINLSGRFSGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFPALMAQ 353
Db      238  AQFLRLIDEGRIASLATAQCCVDESVKYAGERHAFGRNIGAYQAIQFKIADMENKA 297
Qy      354  YV-----MESMTYLTAGMLDQPGPPDCSIEAAMVKVFSSEAAQCSEALQILGGLGYTR 408
Db      298  HMAVGVGRDAASRLVAG---EP-----FKKEAATKLYSSTVAVDNAREATQIHGGYGMN 350
Qy      409  DYPERTLRTRILLIFEGTNEILRMVIA 437
Db      351  EYPAVMWRDLSKILEIGTSETSEVQRMILIA 379

RESULT 9

```

```

US-10-156-761-14127
; Sequence 14127, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14127
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14127

```

```

Query Match
Best Local Similarity 17.6%; Score 556; DB 15; Length 383;
Matches 134; Conservative 84; Mismatches 151; Indels 18; Gaps 9;
Qy      63  ELNEINQFLGPVEKFFTE-EVDSRKI--DOEGKIPDETLEKLSGLGLGLOVPEYGG 119
Db      4  ELSEEQTAVQLAKDFVDREAPNVIAWDRAEEVDSRLVKLGVEVGLGLTVDEEYGG 63
Qy      120  FSNTMYSRLGEIISM-DGSTVTVTLAAHQAIKGLGIIAGTEEQAKYLPKLAGSEHIAAF 178
Db      64  GDHLAYCLVTEELGRDSSVSGIVSVSLGVLAKTIAYWGSSEQKRWLPCLTSGAYVGC 123
Qy      179  CLTEPASGSDAASIRATLSDEKHKHYILNGSKVWITNGLANIFTVFAKTEVDSG 238
Db      124  GLTEPGTGDAGLNTTRAV--RDGDDYVYNGTKMETINGTWADVLLFEARS---TDAP 177
Qy      239  KDKITAFIVERDFGVTNGKPEDKLGIRGNTCEVFHENTKIPVENILGEVCDGFKVAMN 298
Db      178  HKGVSAFLVPTDPTGLTRPIHGKGLRQATAEVLVEDVRVPASAMLAPEKGFSEVAMS 237
Qy      299  ILNSGRFSGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFPALMAQAYVMES 358
Db      238  ALAKGRMSVAACGCVGIAQAALDVAVKYATEREQFGKTIHQLVQELISDIADVDAARL 297
Qy      359  MTYLTAGMLD--QPGFPDCSIEAAMVKVFSSEAAQCSEALQILGGLGYTRDYPYERIL 416
Db      298  LTRVADLVDRQCP-FATASSQA---KLFASEAAVRAANALQVFGGYGYIDEYPAGKLL 353
Qy      417  RTRILLIFEGTNEILRMVIA--ALTGL 441
Db      354  RDARVMTLYEGTSQIKLLIGRALTVG 380

```

```

RESULT 10
US-10-156-761-9454
; Sequence 9454, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262

```

```

; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9454
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9454

Query Match 15.7%; Score 530.5; DB 15; Length 401;
Best Local Similarity 34.8%; Pred. No. 7.3e-43;
Matches 136; Conservative 76; Mismatches 150; Indels 29; Gaps 11;

QY 64 LNEINQ-FLGPYEKFFTEEV--DSRKIDQEGKIPDETLEKLSLGLFLGLOVPEEYGGIGF 120
DB 10 LTDIOQEILSTVRDFVDEKEIIPVATELEHREDEYPOQIVDGLKELGLFLGLMIPEEYGGIGE 69
QY 121 SNTMYSRLEGIISMDG-SITVTAAHQAIGLKGIIA-----GTEOKAKYLPKLASGE 173
DB 70 SLLTVALCEEIARGWMSVGIINTH-----FIVAYMLKQHGTEQOKHFLPRMALGE 122
QY 174 HIAAFCLTEPASGSDAASIRSRATLSSEKXKHYLNGSKVWITNGLANIETVFAKTEVVD 233
DB 123 VRGAFSMSEPALGSDVSAISSKAV--KDGDEVLVNGQKMLNGSSNLVAVLYRSDRGH 180
QY 234 SDGSVKDK-ITAFIVERD--FG-----GVNTRKPEDKLGIRGNTCEVHFENKTKIPVENIL 286
DB 181 PEGTAPHKSWTFLVEKEFGFGEVRGLTIPGKIDKMGYKGVDTTTELIMDGLRVPANRVL 240
QY 287 -GEVGDGFKVAMNLSGRFSMGSVVAGLLKRLIETAYACTRQKFNKRLSEFGLQEK 345
DB 241 GGTTRGFYQMDGVEVGRVNAARGCGVAQRAFLGVSYAQORHTFGKPIAQHQAIOFK 300
QY 346 FALMAOKAYVMESMTYLTAGMLDQPGFPDPCSTEAAMKVFSSFAAQCVSEALQILGLIG 405
DB 301 LAEMATKVEAHAMVNAARKKDSGERND--LEAGMAKYLASEYCKEYCVVEDAFRIHGGYG 358
QY 406 YTRDYPYERILTRILLIFEGTNEILRMVI 436
DB 359 FSKYEYERLYREAPMLLIGEGTAEIQKMI 389

RESULT 11
US-10-156-761-12814
; Sequence 12814, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12814
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12814

Query Match 15.6%; Score 492.5; DB 15; Length 409;
Best Local Similarity 32.8%; Pred. No. 3.8e-39;
Matches 135; Conservative 76; Mismatches 148; Indels 53; Gaps 13;

QY 62 DELNEINQFLGPVEKFFTEEV--DSRKIDQEGKIPDETLEKLSLGLFLGLO-VPEEY-- 115
DB 10 DEQEKVRDWL---HGFAADVIRPAAAEWDERETPPVPIQEAQKVGISLDFYAAQQFD 65
QY 116 -GGIGFSNTMYSRLEGIISMDGSIITVTAAHQAIGLKGIIAGLAGEEOKAKYLPKL---AS 171
DB 66 PTGLGIPMAH-----EELFWGCDAGIALSIVG-TGLAAGVGLANGTEEQIGTWIPQWGDAN 120
QY 172 GEHIAAFCLTEPASGSDAASIRSRATLSSEKXKHYILNGSKVWITNGLANIETVFAKTEV 231
```

```

Db 121 DVKAAFCSSPEAGSDVSSMRVAVYDEAKDEWLVNGTKWTATNGIANVHVVA---V 177
Qy 232 VDSGSKVDKITAFIVERDGGYNGKPKDKLGRSNTCEVHFENTKIPVENILG---- 287
Db 178 VDSELSKGH-ASFIVPNTPGUSQOKFKKHGIRASHTAENVLENARIPGSCLLGKKEK 236
Qy 288 -----EVGDGFKV---AMNLSGRFSMGSVVAGLLKRLIENTAEVACTRKOF 332
Db 237 LDERLARERARAGGERVKNAAMATFEASPAVGAMAVGTARAAYEAELETAKTRQOF 296
Qy 333 NKRLSE-----FGLQEKALMAQKAYVMESMTYLTAGMDQPGFPCDSEAAVMKVFS 387
Db 297 GRFIINQGVAFOLADWRTSIDAARLLVWR-----ASWVYNGKPFATAEGSMKLFAS 350
Qy 388 EAAQCVSEALQTLGLGTRDYPYERILDRTRILLIFEGTNEILLRMYALT 439
Db 351 ETAKKVTAQAVQILGCGGTREYVVERMHRDAIYTFEGTSEIQRLLVIART 402

RESULT 13
US-10-181-319-33
; Sequence 33, Application US/10181319
; Publication No. US20030135032A1
; GENERAL INFORMATION:
; APPLICANT: Lewis, Thomas A.
; APPLICANT: Paszczynski, Andrzej
; APPLICANT: Crawford, Ronald L.
; APPLICANT: Cortese, Marc S.
; APPLICANT: Sebat, Jonathan L.
; TITLE OF INVENTION: Compositions and Methods for Bioremediation
; FILE REFERENCE: IDAHL1940
; CURRENT APPLICATION NUMBER: US/10/181,319
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: PCT/US01/02386
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/177,251
; PRIOR FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Pseudomonas stutzeri
US-10-181-319-33

```

```

Query Match 15.2%; Score 480; DB 12; Length 513;
Best Local Similarity 30.2%; Pred. No. 9.2e-38;
Matches 136; Conservative 94; Mismatches 181; Indels 40; Gaps 12;

Qy 9 RTTAAARACGLVSTANRRLLRTSPVPRAFAPAKELFLGKTKKK-EVPPFPPEVSQ----- 61
Db 70 RATCARRNRQVFSCHLRLLH-HPKRRRARGVPLGOLPPRWOLDPSGRLSRLAGTPE 128
Qy 62 --DEL-----NEINQFLGPVEKFFTEVD--SRKIDQEGKIPDETLEKLA----SLGLEG 108
Db 129 ANDMIDFSLPNEQMLVSTVKRFVENELNPLEDEIERTNAIDSVAGELKQKAREGLWA 188
Qy 109 LQVPEYGGGLFESNTYSLRGEIISMDGTSVTTLAHOAIG-LKGIILAGTEQKAKYLP 167
Db 189 MHMPQEVGGGLSAVEFCLVNEQIGR----TKDVLARRAFGHVPSILVHCTGQREKYLH 244
Qy 168 KLASGEHIAAFLCTEPASGDAASIRSRATLSBDKKHYILNGSKVMTNGGLANITVFA 227
Db 245 AAMRGOTHVSAMSEPEAGSDANGIRT--AVKRDGCEWILNGSKHFSADIASAYIVTA 302
Qy 228 KTEVDSGSKVDKITAFIVERDGGYNGKPKDKLGRSNTCEVHFENTKIPVENILG 287
Db 303 RSE-----EGISCFLVDRDTPGLGPLOEQMMHGRGTHQHGLFFDTCRIAPQQLG 353
Qy 288 EVGDGFKVAMNLSGRES-MGSVAVAGLLKRLIENTAEVACTRKOFNKLSEFGIQRKF 346
Db 354 EPCRGRSLVLGHLNVARLAYVARAVGAMASKLLEMSVDPAKORSQFGAPIGSFQMVQKML 413

```

```

Qy 347 ALMAQKAYVMESMTYLTAGMDQPGFPCDSEAAVMKVFSSEAAHQCVCSEALQTLGLGY 406
Db 414 ADMQCEIYGARMMLLNAWEIDQG--RDVREKVSMLKLFASEMLGRVADSAVOIFGGMGY 471
Qy 407 TRDYPYERILDRTRILLIFEGTNEILLRMYIA 437
Db 472 CTELPYERYRDARVFLYDGTSETHRIMIA 502

```

RESULT 14

```

US-10-205-823-4
; Sequence 4, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-4

```

```

Query Match 15.2%; Score 479; DB 15; Length 415;
Best Local Similarity 30.9%; Pred. No. 8.1e-38;
Matches 137; Conservative 76; Mismatches 195; Indels 36; Gaps 9;

Qy 2 SCCGLFLRTTAAARACGLVSTANRRLLRTSPVPRAFAPAKELFLGKTKKKVFPPEVSQ 61
Db 4 SCRRFRGARGLCLPGRLVLVQTGHRSLTSCIDPSMGLNEE-----QKE---FQKVP 53
Qy 62 DELNEINFLGPVEKFFTEEV--DSRKIDQEGKIPDETLEKLSGLSLGFLQVPEYGG 119
Db 54 D-----FAAREMAPNNAEWDQKLELFPDVMRKAALGFGGVVQTDVGGSG 99
Qy 120 FSNMTYSLRGEIISMD-GSITVTLLAAHQAIKGLGIILAGTEQKAKYLPKLASGEHTA 178
Db 100 LSRDLTSVIFEALATGCTSTTAYISIHNCAMW-IDSEGNEEQRRHKFCPLCTMEKFASY 158
Qy 179 CLTEPASGDAASIRSRATLSBDKKHYILNGSKVMTNGGLANITVFAKTEVVDSDG 238
Db 159 CLTEPGSGDAASLITSAKKQD--HYILNGSKAFISGAGESDIYVVMCR-----GGPG 211
Qy 239 KDKITAFIVERDGGYNGKPKDKLGRSNTCEVHFENTKIPVENILGEVGDGFKVAMN 298
Db 212 PKGISCIYVEKGTPLSFGKKEKVGWNSQPTRAVIFEDCAVPVANNRIGSEGQGLIAVR 271

```

306 INQILMAKYVASKAAVDVAGKAVOLLGANGCHADVAVERYRDAKIMELIESTQIHEIQ 365

Db 306 INQILMAKYVASKAAVDVAGKAVOLLGANGCHADVAVERYRDAKIMELIESTQIHEIQ 365
QY 436 IAL 438
Db 366 IAW 368

Search completed: September 3, 2003, 13:39:34
Job time : 64 secs

QY 299 ILNSGREGSMGVAGLEKRLIEMTAETACRKNRKLSEFGLIQKFAALMAQKAYVMS 358
Db 272 GLNGRINIASCSLGAHAHVILTRDLNVRKQFGEPLASNOYLOFTLADMATRLVAARL 331
QY 359 MYTILTAGMLDQPGPPDSIEAAMVKVSSAAWOCVSEALQILGLGGLYTRDYPERILRD 418
Db 332 MVRNAVAL-OEERKDAVALCSMAKLFATDECAICNOLMGGYGLKDYAVQVVRD 390
QY 419 TRILLIPEGTNEILRMVIALTGLQ 442
Db 391 SRVHILEGSNEVMRILISRLQ 414

RESULT 15
US-10-238-075-951
; Sequence 951, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated from a microorganism, and biological uses of these polynucleotides and of their derivatives
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; PRIOR FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 951
; LENGTH: 376
; TYPE: PRY
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (164)..(164)
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (254)..(254)
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (285)..(285)
; OTHER INFORMATION: unknown
US-10-238-075-951

Query Match 14.4%; Score 452.5; DB 12; Length 376;
Best Local Similarity 31.1%; Pred. No. 2.6e-35;
Matches 113; Conservative 71; Mismatches 152; Indels 27; Gaps 7;
QY 84 SRKIDQEGKIPDETLEKLSGLGLQVPEEYGGGLGFSNTMYSRLGEII-----SMDGSI 138
Db 25 AQNIEROQFIVPDIISCVAGAGYLGASIPQKYGGRGVDSYQLCALHEVMAGVHGSLENLI 84
QY 139 TVTLAAHQATGLKGIIIL--AGTEEQAKAYLPKLASEHIAAFCLETPASGSDAAASIRRA 196
Db 85 TVT-----GMVSTLLQVRGSAQAHAHLPKLATGELIGALIALEPNSGDLVNVETE- 136
QY 197 TLSDEKKHYILNGSKVITNGGLANITFTVFAKTEVVDSDGSKVDKITAFIVERDFGGVTN 256
Db 137 -LQDDGGGWRNLGKKWITLQQTADFFXLI-----HCGNQLATVLIDRNTDGTFTI 186
QY 257 GKPEDKLGIRGNTCEVHFENTKIPVENILGEVGDGPKVAMNT-LSNGRFSMGSVVAGLL 315
Db 187 TPLNDMLGRNLGRLAELHFNDCRLKEDALLGPLTPGVPLAVNFALNEGRETTCAGSLGIC 246
QY 316 KRLIEMTAETACRKNRKLSEFGLIQKFAALMAQKAYVMSYTLTAGMLDQPGFPDC 375
Db 247 QAAVDVAXRYIRQKQKRRLLFSHGIVHFLFATMLTQTXAQLMCF-SAAEYRETLHPAM 305
QY 376 STEAAMVKVFSSEAAWOCVSEALQILGLGGLYTRDYPERILRTRILLIPEGTNEILRMV 435

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:26:07 ; Search time 30 Seconds
(without alignments)
875.835 Million cell updates/sec

Title: US-09-945-326-2
Perfect score: 3153
Sequence: 1 MSGCGFLFRTTAARACRGL.....SQILEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA:*
- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
 - 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
 - 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
 - 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
 - 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pap.*
 - 6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	691.5	21.9	390	4	US-09-252-991A-26058
2	666	21.1	389	4	US-09-328-352-6380
3	650	20.6	384	4	US-09-648-004-6
4	623	19.8	387	4	US-09-328-352-6442
5	618.5	19.6	403	4	US-09-252-991A-27960
6	608	19.3	382	4	US-09-328-352-4850
7	603.5	19.1	419	4	US-09-252-991A-31097
8	596	18.9	394	4	US-09-328-352-6801
9	595.5	18.9	409	4	US-09-364-230-30
10	588.5	18.7	427	4	US-09-364-230-32
11	559.5	17.7	383	4	US-09-328-352-7081
12	514.5	16.3	402	4	US-09-328-352-4281
13	501	15.9	424	4	US-09-252-991A-24653
14	495	15.7	399	4	US-09-328-352-4360
15	495	15.7	464	4	US-09-252-991A-23108
16	472	15.0	282	4	US-09-328-352-6854
17	467.5	14.8	404	4	US-09-328-352-5575
18	465	14.7	408	4	US-09-364-230-28
19	456	14.5	270	4	US-09-252-991A-19574
20	434.5	13.8	863	4	US-09-328-352-4699
21	422.5	13.4	415	4	US-09-252-991A-23105
22	415.5	13.2	333	4	US-09-252-991A-26189
23	394.5	12.5	683	4	US-09-252-991A-22542
24	392	12.4	413	4	US-09-252-991A-21058
25	390	12.4	998	4	US-09-328-352-4516
26	375.5	11.9	825	4	US-09-328-352-5561
27	366.5	11.6	399	4	US-09-328-352-5561

28	360	11.4	659	4	US-09-252-991A-23467	Sequence 23467, A
29	356.5	11.3	401	4	US-09-252-991A-29704	Sequence 29704, A
30	351	11.1	433	4	US-09-252-991A-31231	Sequence 31231, A
31	351	11.1	440	4	US-09-591-095-27	Sequence 27, Appl
32	349.5	11.1	436	4	US-09-591-095-10	Sequence 10, Appl
33	342	10.8	418	4	US-09-252-991A-33098	Sequence 33098, A
34	336.5	10.7	418	4	US-09-328-352-7361	Sequence 7361, Ap
35	328.5	10.4	255	4	US-09-364-230-34	Sequence 34, Appl
36	327.5	10.4	613	4	US-09-328-352-6185	Sequence 6185, Ap
37	319.5	10.1	614	4	US-09-328-352-5130	Sequence 5130, Ap
38	312.5	9.9	595	4	US-09-252-991A-30863	Sequence 30863, A
39	310	9.8	626	4	US-09-651-941-25	Sequence 25, Appl
40	309.5	9.8	361	4	US-09-955-597-25	Sequence 25, Appl
41	309.5	9.8	366	4	US-09-328-352-7718	Sequence 7718, Ap
42	299	9.5	408	4	US-09-252-991A-19861	Sequence 19861, A
43	298.5	9.5	605	4	US-09-252-991A-19796	Sequence 19796, A
44	298.5	9.5	620	4		
45	298.5	9.5				

ALIGNMENTS

RESULT 1

US-09-252-991A-26058
; Sequence 26058, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26058
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26058

Query Match	21.9%	Score	691.5	DB	4	Length	390
Best Local Similarity	37.8%	Pred. No.	1.2e-61				
Matches	147	Conservative	82	Mismatches	121	Indels	39
Gaps	8						
QY	67	INQFLGPVEKFFTEEV--DSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGGLGFSNTM	124				
DB	14	LNLLDLSIRQFVRESILVPHEQVAETDRPEAIARMREMGLFGLSIPAEYGGGLG-----68					
QY	125	YSRLGEITSMDGSIYVT-----LAAHQALGKGIILAGTEEQKAKYLPKLAS	171				
DB	69	-----VTMEPEVSIAPFELGTSAPFRSLLTNNIGSQGIVIDGTBEQKRRVLPRLAS	121				
QY	172	GHIAAFCLTEPASGDAASIRSRATLSKDKKHYTLNGSKVYITNGGLANFTVFAKTEV	231				
DB	122	GELLSFCLTEPDGSDAASLKTAV--RDGEHVYVLTGKRRFTITNAPQAGIYTWART--177					
QY	232	VSDSGSVKDK--ITAFIVERDFGVTKNGKPEDKLGIRGNSNTCEVHFENTKIPVSNILGEV	289				
DB	178	---DPAIRGAGGISAFAVVERGTGSLGKPKRMKGQGAHCTDVFDDCRVPASQLGGV	234				
QY	290	-GDGFKVAMNLTNSGRFSMGSVVAGLLKRLLEMTAEACTRKQNKRLSFEGLIQEKFAL	348				
DB	235	EGVGFKTAMKVLDKRGLHIAAVCVGAERMLDALRYALERKQFGQPIAEFQIQAQMLAD	294				
QY	349	MAQKAYVMESMTYLTAGMLDQPGFPDGSIEAAMVKVFSSEAAAMQCVSEALQILGGLGYTR	408				
DB	295	SKAEAYARCWVIPAARQDEG--RDVGTASCACKLFASEMCGRVADRVAQIFGGAGYIG	352				

QY 409 DYPYERILRDRTRILLIFEETNEILRMVIA 437
 || || || || : || || : || : ||
 nb 353 DYGIERYDVRLEFRIVEGTQIQOLLIA 381

RESULT 2
US-09-328-352-6380
Sequence 6380, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6380
LENGTH: 389
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6380

Query Match	21.1%;	Score 666;	DB 4;	Length 389;	
Best Local Similarity	38.2%;	Pred. No. 4.8e-59;			
Matches 150;	Conservative 79;	Mismatches 148;	Indels 16;	Gaps 8;	
QY	58	EVSDQLNEINQFLGPGVEKFFTEEDSR--KIDQEGKIPDETLEKLKSLGLFGLQVPBEY	115		
DB	5	QMRDE-CMLQQLLSTIRDFVKNELIPRENEVEVKDPIDDIVQOMRELGLFGLTIPEY	63		
QY	116	GGLGFSNTWYSLG--EITSDGSGITVTLLAAHQALGLKGIILAGTTEQAKYLPKLASGEH	174		
DB	64	GGLGITMEEVRAVPFELGQTSAPFRSLGCTNNIGSSAILIDGTEQEKOKYLPFRVSGEI	123		
QY	175	IAAFCLTPASGSDAASIRSRATLSDEKHHYILNGSKVMTITNGLIANTFTVFAKT--EVV	232		
DB	124	IGSFCLTPEPESGSDAASLTSAV--KQDGFYVLNGTKRFTINAPHAATFTVMATNPETK	181		
QY	233	DSGSGVKOKITAFIVERDGGVTNGKPEDKLGIRGNTCEVFHENTKIPVENILGEV-GD	291		
DB	182	GSGG-----ISAFLEAINTPGITLGDIDQMKGQSHTCDVFIEGCRVPASALIGGVEG	236		
QY	292	GFKVAMNIIINSGRFSMGSVVAGLLKRLIEWTAETACTRKQFNKRLSEFLGIEKFAALMAQ	351		
DB	237	GFKTAMKVLGDKRLHIGAYSVGVAERMLNDALNATERKQFGQPIANFOLIQAMLADSKA	296		
QY	352	KAVVMESMTYLTAGMLDQPFDPDCSTEAAMVVFSSAAWQCVCSEALQILGGLYGTRYP	411		
DB	297	EIYAAKCMVLDAARRDNG--ENISTEASCAMKFAFEMCGRVADRCVQIHGGAGYISEYA	354		
QY	412	YERILBDRTRILLIFEGTNEILLRMYIALTGLQHA	444		
DB	355	IERYFDRVRLFRLYEGTTOVQOILIAKNMIREA	387		

RESULT 3
US-09-648-004-6
; Sequence 6, Application US/09648004
; Patent No. 6498242
; GENERAL INFORMATION:
; APPLICANT: CHEN, QIONG
; APPLICANT: THOMAS, STUART
; APPLICANT: NAGARAJAN, VASANTHA
; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND
; TITLE OF INVENTION: INTERMEDIATES
; FILE REFERENCE: CL-1341-A
; CURRENT APPLICATION NUMBER: US/09/648,004
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/252,553
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97

```

; SEQ ID NO 6
; LENGTH: 384
; TYPE: prt
; ORGANISM: Acinetobacter sp.
US-09-648-004-6

```

	Query Match	20.6%;	Score	650;	DB	4;	Length	384;	
	Best Local Similarity	38.6%;	Pred.	No.	2e-57;				
	Matches	154;	Conservative	77;	Mismatches	132;	Indels	36;	Gaps
Qy	61	QDELNE----	I	NQ	LGPVKEFFTEE	VDSRKIDQ	EKIPDETFLEKL	SIGLGLGVLPVEYG	116
Dd	5	QDTNLQVLDMIRQRVDVG-	LIPNEEI----	V	AETDEIPAEIVQQMKELGFLGTIP	EYE	59		
Qy	117	GLGFSTNTMYSLR--	-GEIISMDGSITVTYLA	AHQAI	GLKGII	LATETEOKAKYLPLASGE	173		
Dd	60'	GLGL--TWEEEVYTAFELGR	TSPAPFRSLIGNTNGIS	SGLLIDGSEOKQYFELPLASGE	117				
Qy	174	HIAFCULTEPASGDAA	SIRSRAVLSDSKHYIL	SGKVMTTGLNGLANIFTVFAK--TEV	231				
Dd	118	IIGSFCLTEPDPSGDAA	SLKTAV--KDGDIH	YLINGTKRYITINPHAGVFIVMARTSTEI	175				
Qy	232	VDSGSVKDKTTAFTIVERD	FEGVTNGKPEDKLGR	SGNSTCEVHFENTKPIPVENILGEV-G	290				
Dd	176	KGTGG-----TS	AFIVDSKTPGISLGK	RKKMGQKAHTCDVIPENCRI	PASALIGGVEG	230			
Qy	291	DGFKYAMNILNSRGFS	SMGVAGLLKRLIEMTA	EYACTRKQFNKLSFEGLIQ-----EK	345				
Dd	231	VGFKTAMKVLQGRTH	IAALSVAAGATRMLED	SLOYAVEKFOGQAIFANFOLLIGMLADSK	290				
Qy	346	FALMAQKAYVMESMY	LATGMLDQPFPCDCTEA	AAWVKVFSSEAACQCVCSEALQTLLGGIG	405				
Dd	291	AEIYAACKMWLDAAR	LDRAGQ-----NV	STEASCAMKFATEMCGRVADRQVQHGGAG	343				
Qy	406	YTRDPYPERILDR	TRILLIFEGETNEILR	MYIALTGLOHA	444				
Qy	244	YTSVAIEPEEDVR	IRIVECTTOIOOVI	IARNMIREA	382				

```

RESULT 4
US-09-328-352-6442
; Sequence 6442, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6442
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-6442

```

	Query Match	19.8%;	Score 623;	DB 4;	Length 387;	
	Best Local Similarity	36.8%;	Pred. No. 1.1e-54;			
	Matches 145;	Conservative	79;	Mismatches 130;	Indels 40;	Gaps 10;
Qy	61 QDELNELNQFLGPVERKFFTEEV----	DSRKIDQEGKIPDETLEKLKSLGSLGVLQPVEPYGG	117			
	: :	: : :				
Dd	9 QETLNQLVDMI-----	RQFVEGVLIIPHENEVAETDELPDIVEQMKGALGLGTIPEEYEG	64			
	: :	: : :				
Qy	118 LGSFTNMYSRIGETIIISMDGSITVT-----	LAAHQAIGLKGIILACTEQRAK	164			
	: :	: : :				
Dd	65 LG-----	LTMEEEVVAFELGRTSAPFRSLICNTNGIGSSGLIIDGTEAQKSF	112			
	: :	: : :				
Qy	165 YLPKLASGEHIAFCULTEPAGSGDAASIRATLSDDKKHYTLINGSKWVIITNGGLANIFT	224				
		:				
Dd	113 FLPLRARGEIVISSFCULTEPDAGSDAASLKTSAY--	KDGDFYLINGFKRTFNAPHAGVET	170			
		:				

Qy	225	VFAKTEVSDG	SVKDKITAF	IVRDFG	VGTNGR	PKEDK	LIRG	SNTCEV	HFENTK	IPVEN	284		
		: :	: :	: :	: :	: :	: :	: :	: :	: :			
Db	171	VMARTN	-FDIKGA--	SGISAF	IVDSQT	PGISLGR	KDKKMG	QGVHT	CDVIFEM	CRIPASA	227		
Qy	285	ILGEV	-GDCFKV	AMILNS	GRFSGS	VVAGLL	KRLIEM	TAIEY	ACTRKQ	NRLSEF	GLQ	343	
		:::			:::		:::		:::				
Db	228	LIGVEGV	GFKTAM	VLDKGR	LUHTAALS	VSGAAT	RLDDSL	NYAIE	TKRFG	QOPIAE	FQLIQ	287	
Qy	344	EKFALMAQ	KAYV	WESNTV	LTAGML	DQGF	PCST	EAAMV	KVESSE	AAWCQV	SEALQ	IGG	403
		: :	: :	: :	: :	: :	: :	: :	: :	: :	: :		
Db	288	AMLADSKAE	IIYA	KCMV	-LDAAR	LRDAG	-QNV	STEAS	CAKMFAT	EMCGR	VADRCV	QIHGG	345
Qy	404	LCYTRDYP	PERILL	TRILL	FEGT	NEIL	MYIA					437	
		: :	: :	: :	: :	: :	: :	: :	: :	: :	: :		
Db	346	AGYISEY	AIERFY	RDRV	RLFYEG	TQTQ	QOVI	IA				379	

```

RESULT 5
US-09-252-991A-27960
; Sequence 27960, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27960
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27960

```

Query Match	19.6%;	Score	618.5;	DB	4;	Length	403;
Best Local Similarity	37.5%;	pred. No.	3.5e-54;				
Matches	149;	Conservative	71;	Mismatches	148;	Indels	29;
Gaps							
QY	52	EVFPFFPEVSODELNEINQFPGPVEKFFTEBVDGR--KIDQEGKIPDETLEKLSGLGFL	109				
Db	15	ESMTYPSLNFALGETIDMLRDQVRGVFAAELOPRAAQIDODNPFMDWRKFKFEGMLLGI	74				
QY	110	QVPEEYGGGLFSNTMYSRLCEIIS--MDGSIITVTLLAAHQAIGLAGIILACITPEEQKAKYLPK	168				
Db	75	TVDEEYGGSGALYLAHAVVMEETSRASVALSYGAHSNLCVNOIKRNGNAEQARYLPA	134				
QY	169	LASGEHIAAFLCTEPASGSDAASIRSRATLSEDKHYIILNGSKWITNGSLANIFTVF	228				
Db	135	LVSGEHIGALAMEPNAGSDVSMKLADRVGDR--FVLNGSKWITNGTDPDAHTYVVIK	192				
QY	229	TEVVDSDGSKVKK-----ITAFIVERDFGVGVTNCKPEDKLGIRGSIINCEVHFENKTKTPVEN	284				
Db	193	TD-----DKGAHGITAFIVERDWKFGSRGPKLDKLOWGSIINCELIQDQVEEEN	244				
QY	285	ILGVECDGPKVAMTILNSGRFSWGSVAGLLKRLIEMTAEYACTRQFNKRLSEFGLIOE	344				
Db	245	VLCVANGGVKVLUMSLGDIYERVVSLGGPVGIMQACMDVVVPYITHRRQFGSIOGEFOLVG	304				
QY	345	KF-----ALMAQKAVYMESMTYLTAGMLDOPGPDCEITFAAMVKYFSSAANQCVSEALQ	399				
Db	305	KVADMYTALNASRAVL-----YAVAAACDRG--ETTRDKAAGVILYSAERATOMALDAIQ	357				
QY	400	ILGGILGYTRDYPYERILADTRILLIFEGTNEILRWYI	436				
Db	358	ILGGNGYINERTGTLLRDLAKLEYTGAAGTSETIRMLI	394				

RESULT 6

Query Match 19.18; Score 603.5; DB 4; Length 419;

US-09-328-352-4850
; Sequence 4850, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4850
; LENGTH: 382
; TYPE: prt
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4850

[illegible]

RESULT 7
US-09-252-991A-31097
; Sequence 31097, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074.788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094.190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31097
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31097

Best Local Similarity 38.7%; Pred. No. 1.3e-52;
Matches 136; Conservative 67; Mismatches 139; Indels 9; Gaps 6;

Qy	140	DOECKIPDEYLEKLSIGLFGLOVPEYGGIGLFSNMTYS - RLGESITMSDGSITVTTLAAHQ	140
Db	74	DREHRFAEATREMAEDLUGFLCMLVPEWGGQAOTGLAYAMALEEATAAGDGACSTIMSVHN	133
Qy	147	AIGLKGILAGTEQKAKYLPKLIASGEHIAAFLCTEPAGSGDAASIRSRATLSEDKHYI	206
Db	134	SVGCMPHKTGSAOBKRFRLPLAQSMLGAFALTEPOAGSDASFLLKTRA - RRDGDHIV	191
Qy	207	LNGSKVITNGLLANIETVFAKTEVVDSDGSVVKDITAFIVERDFGGVTNGKPEDKLGIR	266
Db	192	LNGAKQFITSQSHAGWIVFA -- VTDPPDAG - KRGISAFIVPTDTPGYEVVRIEDKLGQH	247
Qy	267	GSNTCEVHFENTKIPVENILGVEGDGFKVAMNITLNSGRFSMGSVVAGLLKLLEMTAYA	326
Db	248	ASDTCOLAFNDLRIPATRLRIGEAGEGYRIALANLEGGRIAGIAQAQVGMARAAFEAARDYA	307
Qy	327	CTRKFQNRKLSEFGLIOEKALMAQKAYVNESMTYLTAGMLDQPPGPDGCSIEAAMVKVFS	386
Db	308	HERETFGPIIEHQAVAFRLADMATRIAVARQMVH - HAASLREAGLP - CLTEASMAKLFA	365
Qy	387	SPAQWCQVSEALQILGGLYGTYTRDYPYPERILLDTRILLIPEGTNEILLMYIA	437
Db	366	SEMAEVCSSAIIQTGGYGYLKDPFVERIYRDRVRQYIYEGTSDVQRLLYA	416

```

RESULT 8
; US-09-328-352-6801
; Sequence 6801, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6801
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-6801

```

Query Match	18.9%	Score 596;	DB 4;	Length 394;
Best Local Similarity	37.5%;	Pred. No. 6.6e-52;		
Mismatches	73;	Mismatches 138;	Indels 22;	Gaps 9;
Conservative	140;	Conservative		

74	QY	VEKFFTEVD--SRKIDQEGKIDPDTLEKLSGLFGLQVPEEYGGGLGFSNTMYSRLGEI	131
25	Db	VAAECAKEIAPIAQQDDHNFHAWKFKFGDMLGLMTVSEYGG---ANWGY--LAHI	79
132	QY	ISM-----DGSITVTLLAAHQALICGLIILACTEQKAKYLPKLASGEHIAFCLTEPAS	185
80	Db	IAMQEISRASAAIGLSYGAHSLCVNQINRNGNEQKQKYLPLKLIISVYVGALAMSEPN	139
186	QY	GSDAASITRSATLSEDKKHVILNGSKVWITNGGLANIEVFAKTEVWSDGSKVDKIATF	245
140	Db	GSDVVSMKLRAEQKGD--HFLVNGSKWITNGSDADVLVYAKT---DPQAGPKG-MTAF	193
246	QY	IVERDFGVTNGKPEDKLGIRGNSNTECVHFENTKIEVENILGEVDGDFKVMANILNSGRF	305
194	Db	LIEKGMKGFSGHNLDLKLGRGNSNTYPLFFPDNYEVAENVLGSGVNGVYLMGSLDYERA	253
306	QY	SMGSVVACLLKRLTEMTAEACTRKQENKRLSBFGLTIOEKFALMAQKAYVMSEMTYLATG	365
254	Db	VLSAGPLGIMADCLDVVPIYLHQREQFGQALGEFQLMOCKLADMYSTWLACKALVYAVGA	313
366	QY	MLDQPGFPDCSI--EAMVKVFSSEAAQCVSEALQILGGLGYTRDPYPYERILRTRILL	423
314	Db	ACDKADH--DRSLURDAASAILYAAAEKATWAGASIAOTLGGNGYNEFPAGLWRDKLYE	372

424 IFEGTNEILURMYI 436
 | | | | |
 373 IGAGTSEIRMLI 385
 Db

RESULT 9
 US-09-364-230-30
 ; Sequence 30, Application US/09364230
 ; Patent No. 6348339
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Hitz, William D.
 ; APPLICANT: Kinney, Anthony J.
 ; APPLICANT: Rafalski, J. Antoni
 ; TITLE OF INVENTION: Enzymes Involved in Degradation
 ; FILE REFERENCE: BB-1178
 ; CURRENT APPLICATION NUMBER: US/09/364,230
 ; CURRENT FILING DATE: 1999-07-29
 ; EARLIER APPLICATION NUMBER: 60/094,990
 ; EARLIER FILING DATE: July 31, 1998
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 30
 ; LENGTH: 409
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-09-364-230-30

[illegible]

```

RESULT# 10
US-09-364-230-32
; Sequence 32, Application US/09364230
; Patent No. 6348339
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Enzymes Involved in Degradation of Branched-Chain Amino Acids
; FILE REFERENCE: BB-1178

```

; CURRENT APPLICATION NUMBER: US/09/364,230
 ; CURRENT FILING DATE: 1999-07-29
 ; EARLIER APPLICATION NUMBER: 60/094,990
 ; EARLIER FILING DATE: July 31, 1998
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 32
 ; LENGTH: 427
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 US-09-364-230-32

Query Match 18.7%; Score 588.5; DB 4; Length 427;
 Best Local Similarity 37.6%; Pred. No. 4.4e-51;
 Matches 143; Conservative 66; Mismatches 144; Indels 27; Gaps 10;

```

QY 69 QFLGPVEKFEETEV--DSRKIDQEGKIPDETLEKLSLG---LFLQVPEEYGGIGFSNT 123
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 55 QKESVAQPATENIAPHASKIDQTYFPKE-VNLWKSMEGFNLLGITAPEYVGGGLGLGYL 113

QY 124 MYS-RLGEIISMDGSTITVTLAAHQIAGLKIILACTEOKAKYLPKLAGSEHIAFCLTE 182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 114 YHCIAEETSRASGVSGLSYGSAHNLINQLVRNGSPAQKEKYLPLKLSGDHVGALAMSE 173

QY 183 PASGSDAASIRSRAITSEDKKHYILNGSKVMTINGLANIFTVFAKTEV-VDSGSKVDK 241
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 174 PMSGDVVSMCKADRVDG--GYVLNGKMMCTNGPVAQTLVYVAKTDITAGSKG----- 226

QY 242 ITAFIVDFDGFVGTNGKPEDKLGINGSNTCEVHFENTKIPVENTILGEVGDGFKVAMNIN 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 227 ITAFIEKGMPCFNFAOKLDLGRGSDTCELVFENCVPDENILGKGGVYVMSGLD 286

QY 302 SGRESMGVSVAGLKLRLIEMTAETACTKQFNKRLSEFGLIOEKFA-----LMAQKAYVM 356
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 287 LERLVLAAGPLGIMQACLDVLPVVRQEQGRIGEFQFTQGIADMYTSLQSSRSYVY 346

QY 357 ESMYTLTAGLDQGPDPDCEAAWVSEAAWQCVSEALQILGLGTYTRDYPYBRIL 416
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 347 SVARDCNGKVDP---KDC-----AGAILCAERATQVALQIOLGNGVYNEYPTGRLL 399

QY 417 RDRTRILLIFETNEILRWYI 436
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 400 RDAKLYEIGACTSBIIRMI 419
  
```

RESULT 11

US-09-328-352-7081
 ; Sequence 7081, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 7081

LENGTH: 383

TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-7081

Query Match 17.7%; Score 559.5; DB 4; Length 383;
 Best Local Similarity 36.9%; Pred. No. 3.3e-48;
 Matches 130; Conservative 74; Mismatches 137; Indels 11; Gaps 7;

```

QY 88 DQSGKIPDETLEKLSGLFLQVPEEYGGIGFSN-TMYSRLGEIISMDGSTITVTLAAHQ 146
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35 DKTHEPQALKDLGALGAMGTVPDEWGGAGMDYVSLVLAETEEAAGDAISTIVSQN 94

QY 147 AICLKGIIILA-GTEEOKAKYLPKLAGSEHIAFCLTEPASGSDAASIRSRAITSEDKKHY 205
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

RESULT 13

US-09-252-991A-24653

; Sequence 24653, Application US/09252991A

```

DB 95 SL-ICGITLAYSEQOKOTYLPKFASEGNLCCFCLTEPHVGSASAILCKA--ERGDHW 151

QY 206 ILNGSKVMTINGLANIFTVFAKTEVVDSDGSKVDKITAFIVERDFGVTNGKPEDKLG 265
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 152 VLVNGKOFITSGKNAQVALFAVTD---KOAGKKGISCFVPTNTOGYLVTRIEDKMGQ 207
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 RGSNTCEVHFENTKIPVENILGEVGDGFKVAMNINLSGRFSMGSVAVGLKLRLIEMTAET 325
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 208 HASDTATITLEDRCRIPLENLVGQEGYKIALSALAAGRIGIAAQSVMARAAPDAVOY 267
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 ACTRKQNKRLSEFGLIOEKPALMAQKAYVMESMTYLTAGMLDQGPDPDCSIEAAMVKVF 385
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 268 ANERAKFGVELVQHAQVGFRLADMATQIEAAHQLV-LHAATLKDAGLP-CLKESAMKLP 325
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 SSEAAWQCVSEALQILGLGTYTRDYPYERITLDRTRILLIFECTNEILRWYIA 437
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 326 ASTMAERVCSDAQIHHGYGVSDFPVERIYRDRVRSQIYEGASDIQRLVIA 377
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

RESULT 12

US-09-328-352-4281

; Sequence 4281, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 4281

; LENGTH: 402

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-4281

Query Match 16.3%; Score 514.5; DB 4; Length 402;

Best Local Similarity 35.4%; Pred. No. 1.4e-43;

Matches 140; Conservative 64; Mismatches 167; Indels 25; Gaps 12;

```

QY 55 PFPEVSDQDEL--NEINQFLGPVEKFEETEV--DSRKIDQEGKIPDETLEKLSGLFLQ 110
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20 PIPNMLPRHLFNEEHEAFRETVRKFEYEVVNPNIKEKQHVDRDLNWKAGSLGLCTT 79

QY 111 VPEEYGGIGFSNTMYSRL---GEIISMDGSTITVTLAAHQIAGLKIILAGTEQKAKYLP 167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 80 MPEQYGGSGV-DRLYSMILIEEQAYAMDSSTGFSL--HSDIVANYINNFGNEQKQWLP 136

QY 168 KLASGEHIAAFLCTEPASGSDAASIRSRATLSSEDKKHYILNGSKVMTINGLANIFTVFA 227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 137 KMATGETVTAIATMEPTGSDLOAVRTTAVLDGDE--YVINGSKIFITNGVLCDMAIVVC 194

QY 228 KTEVVDSDGSKVDKITAFIVERDFGVTNGKPEDKLGIRGNTCEVHFENTKIPVENILG 287
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 195 KTGNSDK-GSA--NLSLIIVEADRAGETKGPLNKGKMGODTCELFDFNVVRPKENLLG 251

QY 288 EVGQGFVKAMNINLSGRFSMGSVAVGLKLRLIEMTAETACTKQFNKRLSEFGLIOEKFA 347
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 252 MEGMGFIMLMKELAWERMLVAICQAGAKAAFAHTVOYTKDKRAGFKPIGAFO--NTRFK 309

QY 348 LMAQKAYVMESMTYLTAGM---LDQPGPDCSIEAAMVKVFSEAAWQCVSEALQILGG 403
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 310 LAELRTEIDFTRTYLDRCMELQLDE---KLSVEAAAAAAYKISDMFSKVVDDECLQHHG 365

QY 404 LGYTRDYPYERILLDRTRILLIFECTNEILRWYIALT 439
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 366 YGYMWEPIARAYIDHRANRIYAGTNEIMKELISRT 401
  
```

Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24653
LENGTH: 424
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24653
Query Match 15.9%; Score 501; DB 4; Length 424;
Best Local Similarity 33.8%; Pred. No. 3.6e-42;
Matches 127; Conservative 69; Mismatches 158; Indels 22; Gaps 8;
QY 70 FLGPVEKFFTEV--DSRKIDQEGKIPDETLEKLSLGLFGLQVPEYGGGFGFNTMYSR 127
DB 60 FRDSVRKFLQEAQVPHQWKGHVDRLNWKAGEAGMCLCSHLPEYGGMG-ADFLYSA 118
QY 128 LG-ETISMDGSTVTLAAHQALGLKGIILAGTEQOKAKYLPKLAGSEHIAAFCLTEPASG 186
DB 119 VVIEETGRAGLTGIGFSLHSDIAAPYILHYSEALKRKLPLKLVSGEMVAIAIATEPGAG 178
QY 187 SDAASIRATLSSEDKKHYILNGSKVWITNGGLANITVFATKTEVVDSDGSKDKITAFI 246
DB 179 SLOGVKITAVLDGE--YVINGSKTFTINGFLADLVIVAKT---DPKAGAKG-ISLFV 232
QY 247 VERDFGGVNGPKDLGIRGNTCEVHFENTKIPVENILGVDGFKVAMNLSGRFS 306
DB 233 VEAGTGFSGKRLEKVGKMAQDTSELFQDVRIPRENLLGKDGOGFIYLMQELPQERLT 292
QY 307 MGSVAGLLKRLIEMTAETACRKFQFNKRLSEFGLIQKFLMAQ-----RAYVMESMTY 361
DB 293 VGIGALASAEALOWTLDTYREKAFGKSVADQFONTRFKLAEMATEIQVGRVFDRCMEQ 352
QY 362 LTAGMLDQPGFPCDSTEAMVVFSEAAWOCVSEALQILGLGYTRDYPYERILDRRI 421
DB 353 HLQCKLDVP-----TAAMCKYWTITDLOCKVLDCEVQJHGGYGFEMWEPYVARAWADRV 405
QY 422 LLIFEGTNEILRMVIA 437
DB 406 QRIYAGTNEIMKEIIA 421
RESULT 14
US-09-328-352-4360
Sequence 4360, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4360
LENGTH: 399
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-4360
Query Match 15.7%; Score 495; DB 4; Length 399;
Best Local Similarity 32.3%; Pred. No. 1.3e-41;
Matches 124; Conservative 70; Mismatches 162; Indels 28; Gaps 10;

QY 66 EINQFLGPVEKFFTEVDSR--KIDQEGKIPDETLEKLSLGLFGLQVPEYGGGFGFNT 123
DB 25 DLELFRDNFKRPMNEHIAPHYDQWREGIMPRVSWSQLGENGFLCVDVPEYGGYGP-T 83
QY 124 MYS-RIGEIIISMDG--SITVTILAAHQALGLKGIILAGTEQOKAKYLPKLAGSEHIAAFCL 180
DB 84 YYSMLVSEASARAGFCALSTAISSHEIAAPYILHIGTEEQKQYWLPMKMTGVEVGAIGM 143
QY 181 TEPASGSDAASIRATLSSEDKKHYILNGSKVWITNGGLANITVFATKTEVVDSDGSKVD 240
DB 144 TEPGAGSDLOSMRTSAILQDD--HYLLNGSKTFTINGFLADLVIVAKT-----DPQARA 196
QY 241 K-ITAFIVERDFGGVNGPKDLGIRGNTCEVHFENTKIPVENILGVDGFKVAMNI 299
DB 197 KGVSIILLVDTHLEGFKKGTNLDKIGLHSQDTSELFQDVRIPRENLLGKDGOGFIYLMQEL 256
QY 300 LNSGRFSMGSVAGLLKRLIEMTAETACRKFQFNKRLSEFGLIQKFLMAQRAY 354
DB 257 LPRETAIASTAVGAIRGAIDLATAVYKERHAFGQPISQFQNTFRFLAQAKIDELATTAF 316
QY 355 VMESMTYLTAGMLDQPGFPCDSTE-AMVVFSEAAWOCVSEALQILGLGYTRDYPYE 413
DB 317 YERNVALYQBGKLD-----VETAAALKSFSTDMQKVNADNLLQLFGGIGYTWTEYPI 368
QY 414 RIIRDTRILLIFEGTNEILRMVIA 437
DB 369 RFFVDARIQRIYGGTNEIMKEIIVA 392
RESULT 15
US-09-252-991A-33108
Sequence 33108, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 33108
LENGTH: 464
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33108
Query Match 15.7%; Score 495; DB 4; Length 464;
Best Local Similarity 32.9%; Pred. No. 1.7e-41;
Matches 144; Conservative 77; Mismatches 179; Indels 38; Gaps 13;
QY 12 AAARACRGLVSTANRRLLRTSPVY-RAFAKELFLGKIKKEVFPFP---EVSODELNE- 66
DB 44 AALRAQRGRCRGA--RRGRETFPSIQRALGAE-----QVPMDFRLTQEQDMLVEA 93
QY 67 INQFLGPVEKFFTEVDSRKKIDQEGKIPDETLEKLSLGLFGLQVPEYGGG--- 119
DB 94 VRSEF---VEKELLPHEDD--VDRADAVSPELAAQIRGKALAAAGFAFNMPPEVGGGLDY 148
QY 120 FSTMYSLRGETIISMDGSTVTLAAHQALGLKGIILAGTEQOKAKYLPKLAGSEHIAAF 179
DB 149 LSQLALVER--ELSKVSWALHVFVAPRSK-----ILMACTGOLGDLPLPCVOGETDCFA 201
QY 180 LPEPASGSDAASIRATLSSEDKKHYILNGSKVWITNGGLANITVFATKTEVVDSDGSKV 239
DB 202 LPEPGAGSDANSIKTRAV--ROGDAFVINGSKHFIHSHAGHADFAIVFAVTDYSYEHNGRKR 259
QY 240 DKITAFIVERDFGGVNGPKDLGIRGNTCEVHFENTKIPVENILGVDGFKVAMNI 299

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
OM protein - protein search, using sw model
Run on: September 3, 2003, 13:24:27 ; Search time 108 seconds
(without alignments)
1483.802 Million cell updates/sec

Title: US-09-945-326-2
Perfect score: 3153
Sequence: 1 MSCGGLFLRTTAARACRL.....SQILEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62
Gapop 10.0 , Gapext -0.5
Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3153	100.0	621	4 Q9H845	Q9H845 homo sapien
2	3149	99.9	621	4 Q8WXX3	Q8WXX3 homo sapien
3	2704	85.8	625	11 Q8JZN5	Q8JZN5 mus musculus
4	2703	85.7	625	11 Q8BK76	Q8BK76 mus musculus
5	2700	85.6	625	11 Q8C0B5	Q8C0B5 mus musculus
6	2526	80.1	498	4 Q9H9W4	Q9H9W4 homo sapien
7	1318	41.8	655	6 Q8HXV7	Q8HXV7 macaca fasc
8	1308	41.5	633	4 Q8WU10	Q8WU10 homo sapien
9	1295.5	41.1	655	5 Q9V8T1	Q9V8T1 drosophila
10	1210.5	38.4	613	5 Q19057	Q19057 caenorhabdi
11	954	30.3	188	4 Q9BUX5	Q9BUX5 homo sapien
12	821	26.0	594	16 Q9K784	Q9K784 bacillus ha
13	815.5	25.9	594	16 Q82176	Q82176 bacillus su
14	789.5	25.0	594	16 Q8ENT4	Q8ENT4 oceanobacil
15	734	23.3	169	11 Q9CZC5	Q9CZC5 mus musculus
16	719.5	22.8	382	16 Q8EM49	Q8EM49 oceanobacil

17	709.5	22.5	381	16 Q8R674	Q8R674 fusobacteri
18	699	22.2	382	16 Q9AB07	Q9AB07 caulobacter
19	699	22.2	565	16 Q8ES29	Q8ES29 oceanobacil
20	695.5	22.1	380	16 Q8RC01	Q8RC01 thermoaer
21	691.5	21.9	384	16 Q9I391	Q9I391 pseudomonas
22	690	21.9	379	16 Q9K6D0	Q9K6D0 bacillus ha
23	688	21.8	380	16 Q8FUX4	Q8FUX4 brucella su
24	682	21.6	381	16 Q8YDG3	Q8YDG3 brucella me
25	681	21.6	380	2 Q8VOK1	Q8VOK1 brucella ab
26	678	21.5	381	16 Q9A7C6	Q9A7C6 caulobacter
27	672.5	21.3	375	2 Q9AHX9	Q9AHX9 pseudomonas
28	667	21.2	379	16 Q8EM50	Q8EM50 oceanobacil
29	659	20.9	384	16 Q8FX67	Q8FX67 brucella su
30	658	20.9	380	16 Q9K6D1	Q9K6D1 bacillus ha
31	655.5	20.8	389	2 Q8KZQ2	Q8KZQ2 butyrivibri
32	654.5	20.8	379	2 Q8RM16	Q8RM16 clostridium
33	652	20.7	387	16 Q9RUX5	Q9RUX5 deinococcus
34	651.5	20.7	384	2 Q93772	Q93772 acinetobacter
35	650.5	20.6	636	16 Q8RSV8	Q8RSV8 fusobacteri
36	650	20.6	384	2 Q9F7E6	Q9F7E6 acinetobact
37	650	20.6	384	16 Q8ESF1	Q8ESF1 oceanobacil
38	648	20.6	379	2 P97088	P97088 clostridium
39	642.5	20.4	385	16 Q8XR24	Q8XR24 ralstonia s
40	642	20.4	658	16 Q9KYS2	Q9KYS2 streptomyce
41	633.5	20.1	379	16 Q8X124	Q8X124 clostridium
42	629.5	20.0	381	16 Q9RVV0	Q9RVV0 deinococcus
43	628.5	19.9	385	16 Q9RIQ5	Q9RIQ5 streptomyce
44	628	19.9	456	16 Q8YB77	Q8YB77 brucella me
45	626	19.9	380	16 Q8UAA0	Q8UAA0 agrobacteri

ALIGNMENTS

RESULT 1

ID	Q9H845	PRELIMINARY;	PRT;	621 AA.
Q9H845				
AC	Q9H845			
DT	01-MAR-2001 (TReMBLrel. 16, Created)			
DT	01-MAR-2001 (TReMBLrel. 16, Last sequence update)			
DT	01-MAR-2003 (TReMBLrel. 23, Last annotation update)			
DE	Hypothetical protein FLJ13950.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuma M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,			
RA	Ninomiya K., Iwayanagi T.;			
RT	NEDO human cDNA sequencing project.;			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=Lung, and Uterus;			
RA	Strausberg R.;			
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AK024012; BAB14775.1;			
DR	EMBL; BC013354; AAHL13354.1;			
DR	EMBL; BC007970; AAH07970.1;			
DR	HSSP; Q06319; 1BUC.			
DR	InterPro; IPR006089; Acyl-CoA_dh.			
DR	InterPro; IPR006090; Acyl-CoA_dh_C.			
DR	InterPro; IPR006091; Acyl-CoA_dh_M.			
DR	InterPro; IPR006092; Acyl-CoA_dh_N.			
DR	InterPro; IPR005829; Sug_transporter.			
DR	Pfam; PF00441; Acyl-CoA_dh; 1.			
DR	Pfam; PF02770; Acyl-CoA_dh_M; 1.			

DR	Pfam: PF02771; Acyl-CoA_dh.N: 1.
DR	PROSITE: PS00072; ACYL_COA_DH_1: 1.
DR	PROSITE: PS00073; ACYL_COA_DH_2: 1.
DR	PROSITE: PS00216; SUGAR_TRANSPORT_1: 1.
DR	KW Hypothetical protein.
SQ	SEQUENCE 621 AA; 68760 MW; 064BCE0378877F54 CRC64;
	Query Match 100.0%; Score 3153; DB 4; Length 621;
	Best Local Similarity 100.0%; Pred.No.5.le-214;
	Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MSGCGLFLRTTAAARACRGLVSTANRRLLRTSPPVRAFAKELFGKIKKKKEVFPPPEVS 60
DB	
QY	1 MSGCGLFLRTTAAARACRGLVSTANRRLLRTSPPVRAFAKELFGKIKKKKEVFPPPEVS 60
DB	
QY	61 QDELNEINQFTGPVEKFFTEVDSSRKIDQEGKIPDETLEKLKSLGFLGQVPEEYGGIGF 120
DB	
QY	61 QDELNEINQFTGPVEKFFTEVDSSRKIDQEGKIPDETLEKLKSLGFLGQVPEEYGGIGF 120
DB	
QY	121 SNMTYSRLGEIISMDGSITVTLAHQAIGLKGIIILAGTEQKAKYLPKLASGEHIAAFCL 180
DB	
QY	121 SNMTYSRLGEIISMDGSITVTLAHQAIGLKGIIILAGTEQKAKYLPKLASGEHIAAFCL 180
DB	
QY	181 TEPASGDAASTRSRATLSSEDKHYIINGSKVWTNGGLANIPTFAKTEVVDSDGVKD 240
DB	
QY	181 TEPASGDAASTRSRATLSSEDKHYIINGSKVWTNGGLANIPTFAKTEVVDSDGVKD 240
DB	
QY	241 KITAFIVDFGVGTNGAPEDKLGIRGNTCEVHFENTKIPVENILGEVDGFKVAMNII 300
DB	
QY	241 KITAFIVDFGVGTNGAPEDKLGIRGNTCEVHFENTKIPVENILGEVDGFKVAMNII 300
DB	
QY	301 NSGRFSMGSVVAGLLKRLIEMTAETACTRKQFNKRSLSEFGLIQEKPALMAQAAYVMESMT 360
DB	
QY	301 NSGRFSMGSVVAGLLKRLIEMTAETACTRKQFNKRSLSEFGLIQEKPALMAQAAYVMESMT 360
DB	
QY	361 YLTAGMLDQPGPDCSIKAAMVKVFSSEAQCVCSEALQLLGGIGYTDRDPYERILIIRDTR 420
DB	
QY	361 YLTAGMLDQPGPDCSIKAAMVKVFSSEAQCVCSEALQLLGGIGYTDRDPYERILIIRDTR 420
DB	
QY	421 ILILFEGETNEILLRMVIALTLGLQHAGRILTTTRIHELKQAKYSTVMDTVGRRLDSLGRTVD 480
DB	
QY	421 ILILFEGETNEILLRMVIALTLGLQHAGRILTTTRIHELKQAKYSTVMDTVGRRLDSLGRTVD 480
DB	
QY	481 LGITGNHGWHVPSLADSANKPFENYICFRIVEILLLRFGKTIMEEOLVLKRVANILINL 540
DB	
QY	481 LGITGNHGWHVPSLADSANKPFENYICFRIVEILLLRFGKTIMEEOLVLKRVANILINL 540
DB	
QY	541 YGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQDKYAPENLDQIKK 600
DB	
QY	541 YGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQDKYAPENLDQIKK 600
DB	
QY	601 VSQQILEKRAYICAHPDLRTC 621
DB	
QY	601 VSQQILEKRAYICAHPDLRTC 621
DB	
RESULT 2	
Q8WXX3	PRELIMINARY; PRT; 621 AA.
ID	Q8WXX3
AC	Q8WXX3;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Very-long-chain acyl-CoA dehydrogenase VLCAD.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxId=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Zhang W., Zhang J., Li N., Wan T., Chen T., Zhang M., Cao X.;
RL	Submitted (DEC-2000) to the EMBL/GenBank/DDJB databases.
DR	EMBL: AF327351; AAL56011.1; -

OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC031137; AAH31137.1; -
DR EMBL: BC032213; AAH32213.1; -
DR EMBL: BC033277; AAH33277.1; -
DR InterPro: IPR006089; Acyl-CoA_dh.
DR InterPro: IPR006090; Acyl-CoA_dh.C.
DR InterPro: IPR006091; Acyl-CoA_dh.M.
DR InterPro: IPR006092; Acyl-CoA_dh.N.
DR Pfam: PF00441; Acyl-CoA_dh; 1.
DR Pfam: PF02770; Acyl-CoA_dh.M; 1.
DR Pfam: PF02771; Acyl-CoA_dh.N; 1.
DR PROSITE: PS00072; ACYL_COA_DH.1; 1.
DR PROSITE: PS00073; ACYL_COA_DH.2; 1.
KW Hypothetical protein.
SQ SEQUENCE 625 AA; 68707 MW; 4F06FFBFD82F022 CRC64;

Query Match 85.8%; Score 2704; DB 11; Length 625;
Best Local Similarity 85.7%; Pred. No. 2.6e-182;
Matches 534; Conservative 38; Mismatches 47; Indels 4; Gaps 1;

QY 1 MSGCGLFLR-----TTAARACRGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPF 56
Db 1 MSGCVLLSRGATAAAAAAASRVLRREFTARRRPLHTSLQSCSPAKLEFLGNIEQKGVFPF 60
QY 57 PEVSQDELNEINQPLGVPEKFFTEVDSDRKIDQEGKIPDETLEKLSLGLFGLQVPEYG 116
Db 61 PEVSQDELNEINQPLGVPEKFFTEVDSDRKIDQEGKIPDETLEKLSLGLFGLQVPEYG 120
QY 117 GLGFSNTMYSRLEIISMGSGITVTAAHQAIKGLIILAGTBEQAKYLPKLASGEHTA 176
Db 121 GLGFSNTMYARLGEIISLDSITVTAAHQAIKGLIILVGNDEQAKYLPKLSSGEHTA 180
QY 177 AFCLTEPASGSDAASIRSRATLSLSEKHHYILNGSKVWITNGGLANITFTVFAKTEVVDSDG 236
Db 181 AFCLTEPASGSDAASITQTRATLSLSEKHHYILNGSKVWITNGGLANITFTVFAKTEVVDSDG 240
QY 237 SVKDKITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEGVGGFKVA 296
Db 241 SKTDKMTAFIVERDFGGITNGKPKEDKLGIRGNTCEVHFENTRVPVENILGEGVGGFKVA 300
QY 297 MNILNSGRFSMGSVVAGLLKRLIEMTAETAYACTRKQFNKRLSEFGLIOEKFALMAQKAYVM 356
Db 301 MNILNSGRFSMGSAVAGMLKLLIETAEYACTRKQFNRLNLSFGLIOEKFALMAQKAYVM 360
QY 357 ESMYTLTAGMDQPGPDCSIEAAMVKVFSSEAAQCVCSEALQILGGLGYTRDYPYERIL 416
Db 361 ESMAYLTSGMDQPGPDCSIEAAMVKVFSSEAAQCVCSEALQILGSGYMKDYPYERML 420
QY 417 RDRILLIFEGTNEILRMVIALTGLQAGRIILTRIHELKQAKVSTVMDTVGRRRLDSLG 476
Db 421 RDRILLIFEGTNEILRFLIATGLQAGRIILTSRIKELKSGNVTTVMETIGRKLRLDSLG 480
QY 477 RTVDLGLTGNHGVHPSLADSANKPEENTYCFGRVETLLRFGKTTIMEEOLVLRKRVANI 536
Db 481 RTVDLGLTGLGVHPSLSDSANKLEENVHYFGRVETLLRFRGNIVEEOLVLRKRVANI 540
QY 537 LINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLSQDKYAPENLDE 596
Db 541 LINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLSQDKYAPENLDE 600
QY 597 QIKKVSQOILEKRAYICAHPLDR 619
Db 601 QIKKVSQOILEKRAYICAHPLDR 623

RESULT 4
Q8BK76
ID Q8BK76 PRELIMINARY; PRT; 625 AA.

AC Q8BK76;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Very-long-chain acyl-CoA dehydrogenase VLCAD homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL: AK075984; BAC36096.1; -
SQ SEQUENCE 625 AA; 68722 MW; 29567F60B52E6FEA CRC64;

Query Match 85.7%; Score 2703; DB 11; Length 625;
Best Local Similarity 85.7%; Pred. No. 3e-182;
Matches 534; Conservative 38; Mismatches 47; Indels 4; Gaps 1;

QY 1 MSGCGLFLR-----TTAARACRGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPF 56
Db 1 MSGCVLLSRGATAAAAAAASRVLRREFTARRRPLHTSLQSCSPAKLEFLGNIEQKGVFPF 60
QY 57 PEVSQDELNEINQPLGVPEKFFTEVDSDRKIDQEGKIPDETLEKLSLGLFGLQVPEYG 116
Db 61 PEVSQDELNEINQPLGVPEKFFTEVDSDRKIDQEGKIPDETLEKLSLGLFGLQVPEYG 120
QY 117 GLGFSNTMYSRLEIISMGSGITVTAAHQAIKGLIILAGTBEQAKYLPKLASGEHTA 176
Db 121 GLGFSNTMYARLGEIISLDSITVTAAHQAIKGLIILVGNDEQAKYLPKLSSGEHTA 180
QY 177 AFCLTEPASGSDAASIRSRATLSLSEKHHYILNGSKVWITNGGLANITFTVFAKTEVVDSDG 236
Db 181 AFCLTEPASGSDAASITQTRATLSLSEKHHYILNGSKVWITNGGLANITFTVFAKTEVVDSDG 240
QY 237 SVKDKITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEGVGGFKVA 296
Db 241 SKTDKMTAFIVERDFGGITNGKPKEDKLGIRGNTCEVHFENTRVPVENILGEGVGGFKVA 300
QY 297 MNILNSGRFSMGSVVAGLLKRLIEMTAETAYACTRKQFNKRLSEFGLIOEKFALMAQKAYVM 356
Db 301 MNILNSGRFSMGSAVAGMLKLLIETAEYACTRKQFNRLNLSFGLIOEKFALMAQKAYVM 360
QY 357 ESMYTLTAGMDQPGPDCSIEAAMVKVFSSEAAQCVCSEALQILGGLGYTRDYPYERIL 416
Db 361 ESMAYLTSGMDQPGPDCSIEAAMVKVFSSEAAQCVCSEALQILGSGYMKDYPYERML 420
QY 417 RDRILLIFEGTNEILRMVIALTGLQAGRIILTRIHELKQAKVSTVMDTVGRRRLDSLG 476
Db 421 RDRILLIFEGTNEILRFLIATGLQAGRIILTSRIKELKSGNVTTVMETIGRKLRLDSLG 480
QY 477 RTVDLGLTGNHGVHPSLADSANKPEENTYCFGRVETLLRFGKTTIMEEOLVLRKRVANI 536
Db 481 RTVDLGLTGLGVHPSLSDSANKLEENVHYFGRVETLLRFRGNIVEEOLVLRKRVANI 540
QY 537 LINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLSQDKYAPENLDE 596
Db 541 LINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLSQDKYAPENLDE 600
QY 597 QIKKVSQOILEKRAYICAHPLDR 619
Db 601 QIKKVSQOILEKRAYICAHPLDR 623

RESULT 5
Q8C0B5
ID Q8C0B5 PRELIMINARY; PRT; 625 AA.

AC QH9W4;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein FL12506.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT "NEDO human cDNA sequencing project";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK022568; BAB14104.1;
DR HSSP; Q06319; IBUC.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR InterPro; IPR005829; Sug. transporter.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_L; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 498 AA; 55083 MW; A57B7121B8FB1310 CRC64;

Query Match 80.1%; Score 2526; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 6,7e-170; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 0;

QY 124 MYSRLGEIISMDGSIIVTLAAHQAIKGLIILAGTEEQKAKYLPKLIASGEHIAAFCLETP 183
DB 1 MYSRLGEIISMDGSIIVTLAAHQAIKGLIILAGTEEQKAKYLPKLIASGEHIAAFCLETP 60

QY 184 AGSDAASIRSRATLSDEKHHYILNGSKVWITNGGLANIFTVFAKTEVVDSDGSKVDKIT 243
DB 61 AGSDAASIRSRATLSDEKHHYILNGSKVWITNGGLANIFTVFAKTEVVDSDGSKVDKIT 120

QY 244 AFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNINSG 303
DB 121 AFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNINSG 180

QY 304 RFSMGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFALMAQKAYVNESMTYLT 363
DB 181 RFSMGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFALMAQKAYVNESMTYLT 240

QY 364 AGMLDQPGFDCSTEAAWKVFSSEAAQCSEALQILGGIGYTRDYPYERILDRTRILL 423
DB 241 AGMLDQPGFDCSTEAAWKVFSSEAAQCSEALQILGGIGYTRDYPYERILDRTRILL 300

QY 424 IFECTNEILRMVIALTGLQHAGRIITRIHELKQAKVSTVMDTVGRRLRSLGRVDLGL 483
DB 301 IFECTNEILRMVIALTGLQHAGRIITRIHELKQAKVSTVMDTVGRRLRSLGRVDLGL 360

QY 484 TGNHGVHPSLADSKANKFEENTYCFGRVETLLLRFGKTIIMEQLVKRANILINLYGM 543
DB 361 TGNHGVHPSLADSKANKFEENTYCFGRVETLLLRFGKTIIMEQLVKRANILINLYGM 420

QY 544 TAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFLSLQDLKYPENIDQIKKVSQ 603
DB 421 TAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFLSLQDLKYPENIDQIKKVSQ 480

QY 604 QILEKRAYICAHPLDRTC 621
DB 481 QILEKRAYICAHPLDRTC 498

AC QRC0B5;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Very-long-chain acyl-CoA dehydrogenase VICAD homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
DR EMBL; AK031820; BAC27365.1;
SQ SEQUENCE 625 AA; 68694 MW; 88B0420E6A85D48A CRC64;

Query Match 85.6%; Score 2700; DB 11; Length 625;
Best Local Similarity 85.6%; Pred. No. 4.9e-182;
Matches 533; Conservative 39; Mismatches 47; Indels 4; Gaps 1;

QY 1 MSGCGLEFLR-----TTAARACRGLVSTANRLLRTSPVRAFKEFLGKIKKVEFPF 56
DB 1 MSGCVLLSRGATAAGAARASRVLRFTARRRPLHTSLQCSFAKELFLGNIEQKGVFPF 60

QY 57 PEVSODELNEINQFLGPVFFTEEDVSRKIDQSGKIPDETLEKLSLGLFGLQVPEYG 116
DB 61 PEVSQHELSEINQFVGLPEKFEETEEVDVSRKIDQSGKIPVDTEKLSLGLFGLQVPEYG 120

QY 117 GLGFSNTMYSRLEIISMDGSIIVTLAAHQAIKGLIILAGTEEQKAKYLPKLIASGEHIA 176
DB 121 GLGFSNTMYSRLEIISMDGSIIVTLAAHQAIKGLIILAGTEEQKAKYLPKLIASGEHIA 180

QY 177 AFCLETPASGDAASIRSRATLSDEKHHYILNGSKVWITNGGLANIFTVFAKTEVVDSDG 236
DB 181 AFCLETPASGDAASIRSRATLSDEKHHYILNGSKVWITNGGLANIFTVFAKTEVVDSDG 240

QY 237 SYVKDITAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEVGDGPKVA 296
DB 241 SKTDKMTAFIVERDFGGVTKGPKEDKLGIRGNTCEVHFENTKIPVENILGEVGDGPKVA 300

QY 297 MNILNSGRFSMGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFALMAQKAYVM 356
DB 301 MNILNSGRFSMGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFALMAQKAYVM 360

QY 357 ESMYLTAGMLDQPGFDCSTEAAWKVFSSEAAQCSEALQILGGIGYTRDYPYERIL 416
DB 361 ESMYLTAGMLDQPGFDCSTEAAWKVFSSEAAQCSEALQILGGIGYTRDYPYERIL 420

QY 417 RDRTRILLIEFGTNEILRMVIALTGLQHAGRIITRIHELKQAKVSTVMDTVGRRLRSLG 476
DB 421 RDRTRILLIEFGTNEILRMVIALTGLQHAGRIITRIHELKQAKVSTVMDTVGRRLRSLG 480

QY 477 RTVDLGLTGNHGVHPSLADSKANKFEENTYCFGRVETLLLRFGKTIIMEQLVKRANILINLYGM 536
DB 481 RTVDLGLTGNHGVHPSLADSKANKFEENTYCFGRVETLLLRFGKTIIMEQLVKRANILINLYGM 540

QY 537 LINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFLSLQDLKYPENIDQ 596
DB 541 LINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFLSLQDLKYPENIDQ 600

QY 597 QIKKVSQILEKRAYICAHPLDR 619
DB 601 QIKKVSQILEKRAYICAHPLDR 623

PRELIMINARY; PRT; 498 AA.
RESULT 6
Q9H9W4
ID Q9H9W4

DB	Q9V8T1	PRELIMINARY;	PRT;	655 AA.	
AC	Q9V8T1				
DT	01-MAY-2000	(T-EMBLrel. 13, Created)			
DT	01-MAY-2000	(T-EMBLrel. 13, Last sequence update)			
DT	01-OCT-2002	(T-EMBLrel. 22, Last annotation update)			
DE	CG7461	protein.			
GN	CG7461				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RP	[1]	SEQUENCE FROM N.A.			
RP	STRAIN=BERKELEY;				
RC	MEDLINE=20196006;	PubMed=10731132;			
RA	Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazer V., Zhang Q., Chen L.X.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,				
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Dubin K.J., Evangelista A.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.P.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,				
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,				
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,				
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,				
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,				
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,				
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,				
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,				
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;				
RT	"The genome sequence of Drosophila melanogaster."				
RL	Science 287:2185-2195(2000).				
DR	EMBL: AE003796; AAF57579.1;				
DR	HSSP: Q06319; IBUC.				
DR	FlyBase; FBgn0034432; CG7461.				
DR	InterPro; IPR006089; Acyl-CoA_dh.				
DR	InterPro; IPR006090; Acyl-CoA_dh.				
DR	InterPro; IPR006091; Acyl-CoA_dh.				
DR	InterPro; IPR006092; Acyl-CoA_dh.				
DR	Pfam; PF00441; Acyl-CoA_dh; 1				
DR	Pfam; PF02770; Acyl-CoA_dh; 1				
DR	Pfam; PF02771; Acyl-CoA_dh; 1				
DR	PROSITE; PS00072; ACYL_COA_DH; 1				
DR	PROSITE; PS00073; ACYL_COA_DH; 2				
DR	SEQUENCE 655 AA; 71326 MW; 00916BD701E127CB CRC64;				
Query Match	41.1%;	Score 1295.5;	DB 5;	Length 655;	
Best Local Similarity	45.1%;	Pred. No. 7.8e-83;			
Matches 280;	Conservative 104;	Mismatches 212;	Indels 25;	Gaps 9;	
14 ARACRGIVV-----STANRRLLRTSPVRAFAKELFLGKIKKKEVFPFPPEVSQDELNEI 67					
QY					

DB	Q19057	PRELIMINARY;	PRT;	613 AA.	
AC	Q19057				
DT	01-NOV-1996	(T-EMBLrel. 01, Created)			
DT	01-NOV-1996	(T-EMBLrel. 01, Last sequence update)			
DT	01-MAR-2003	(T-EMBLrel. 23, Last annotation update)			
DE	Hypothetical 66.2 kDa protein.				
GN	E04F6.5				
OS	Caenorhabditis elegans.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;				
OC	Rhabditidae; Peloderinae; Caenorhabditis.				
OX	NCBI_TaxID=6239;				
RP	[1]	SEQUENCE FROM N.A.			
RP	STRAIN=Bristol N2;				
RC	MEDLINE=99069613; PubMed=9851916;				
RA	None;				
RT	"Genome sequence of the nematode C. elegans: a platform for				
RT	investigating biology. The C. elegans Sequencing Consortium."				
RL	Science 282:2012-2018(1998).				
RP	[2]	SEQUENCE FROM N.A.			
RP	STRAIN=Bristol N2;				
RA	Pauley A.;				
RT	"The sequence of C. elegans cosmid E04F6."				
RL	Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				

```
RC STRAIN-Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL EMBL; U28943; (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR HSSP; Q06319; 1BUC.
DR WormPep; E04F6.3; CE01217.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Hypothetical protein.
SQ
  QUERY MATCH 613 AA; 66171 MW; C1F2883ECE34551B CRC64;
  Best Local Similarity 38.4%; Score 1210.5; DB 5; Length 613;
  Matches 272; Conservative 109; Mismatches 205; Indels 37; Gaps 14;

QY 7 FLRTTAAACRGLVSTANRLRTSPVPFAFAKELFLGKIKKEVFFPFVPSQDELNE 66
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 16 FIRLSHSAAK----DAKPKVAADVSP--SFVMNLFKGAQVTDQVFYPLNMTDEQKE 69
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 67 -INQFLGPVEKFTTEEDSRKIDQEGKIPDETLEKLSGLFLQVPEEVGLGFSNTWY 125
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 TCGVMVSPLEKMLVEVDVVKNDSDIPRAVLDFAEGLTFCGLVLPPELESNGNSOM 129
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 126 SRIGEII-SMDSITVTLAAHQAGLIGLIIAGTEQKAKYLPKIASGHEHIAFCLETPA 184
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 ARVAELVAGYDGLGVGVVGAHOSIGYKGLLEGTDAQOKYLPDLATGRKFAAFALTEPT 189
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 185 SGSDAASIRSRATLSBKKHYILNGSKVITNGGLANIEFTVFAKTEVDSDGSVKDKITA 244
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 TCGSDASSVTRAEISADGRHYLVNGKIKWISNGGADVFVFAQTPVQKQADGSKDKMSA 249
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 245 FIVERDFGVTKGPKEDKIGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNINSGR 304
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 FIVERAFGVTSPQEKKGIRGNTTEVHFONLIPVENILGKEGEGFKVAMNINLNGR 309
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 305 FMSGVAVGLLRKRLTEMTAEYACTRKQFNKRLSEGLTOERFALMAQKAYVWESYTLTA 364
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 FGIPACTGAMKHCQKTQVHDHTTRVQFGKKLQEGFNQKLEVMISKLYATESIVYMLS 369
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 365 GMLDQPGPDCSIEAAMVKVFSEAAQCVSALQILGGLGYTRDPYVERILRDRILLI 424
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 SNMDR-GIKYOLEAAGKVLASENAWLVCDDAIQVHGGMGFMRTGLERVLRLRIPRI 428
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 425 FEGTNEILRMVYALTGLQHAGRLITRIHELKQAKVSTVMDTVGRRRLRDSLGRTVDLGLT 484
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 429 FEGANDVURLFTALTGAQHAGK-----HLAEQA--SGVGLIGLAVSRVTG-----GNT 475
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 485 G-NHG-VVHPSIADSSANKFEENTYCFGRVTETLLRLFGKTTIMEQLVLRKRVANILINLYG 542
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 476 GSNFGQVVDASLQDSAKVLDDQIALFGQTVGLLKKHKKGIIDROYEMHRVADAAINIYS 535
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 543 MTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSLQDKYAPENLDEQIKKVS 602
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 536 SAAVLSRATYAIKNKSSADPERKA-TYYVDKAMK-----SSNRFLKDGSGVENASKVAT 591
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 603 QOILEKRAYICA-----HPLD 618
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 592 IESLAKE--VCGNGGLTLQHPVE 612
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Q98UX5
ID Q98UX5 PRELIMINARY; PRT; 188 AA.
AC Q98UX5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to NPD002 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC001817; AA01817.1; -.
DR InterPro; IPR005829; Sug_transporter.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
SQ
  QUERY MATCH 188 AA; 21237 MW; BDB5855A78A6DBB6 CRC64;
  Best Local Similarity 100.0%; Pred. No. 1.4e-59;
  Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 MYIALTGLQHAGRLITRIHELKQAKVSTVMDTVGRRRLRDSLGRTVDLGLTGNHGVHPS 493
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MYIALTGLQHAGRLITRIHELKQAKVSTVMDTVGRRRLRDSLGRTVDLGLTGNHGVHPS 60
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 494 LADSANKFEENTYCFGRVTETLLRLFGKTTIMEQLVLRKRVANILINLYGMTAVLSRASRS 553
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 LADSANKFEENTYCFGRVTETLLRLFGKTTIMEQLVLRKRVANILINLYGMTAVLSRASRS 120
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 554 ITRIGLRNHDHEVLANTFCVEAYLQNLFSLQDKYAPENLDEQIKKVSQOILEKRAYIC 613
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 ITRIGLRNHDHEVLANTFCVEAYLQNLFSLQDKYAPENLDEQIKKVSQOILEKRAYIC 180
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 614 AHPLDRTC 621
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 AHPLDRTC 188
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
Q9K784
ID Q9K784 PRELIMINARY; PRT; 594 AA.
AC Q9K784;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Butyryl-CoA dehydrogenase.
GN BH3486.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RL halodurans and genomic sequence comparison with Bacillus subtilis."
PL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001519; BAB07205.1; -.
DR HSSP; Q06319; 1BUC.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00073; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Complete proteome.
SQ
  QUERY MATCH 594 AA; 65385 MW; D959657767C1B0A9 CRC64;
```

Query Match		26.0%; Score 821; DB 16; Length 594;
Best Local Similarity		33.2%; Pred. No. 2,1e-49;
Matches		207; Conservative 105; Mismatches 210; Indels 102; Gaps 13;
Qy	45	LGKIKKKEVPPPEVSODELNEINQFLGPVEKFT-----EEVDSRKIDQEGKIPDE 96
Db	16	LDEIDADKVF-----TPEDTTEHEMIGKTTEFEVNVPPVLDTEENHQFD-----ISRR 66
Qy	97	TLKLSLGLFLGLOVPEVGGGLFSNTMYSLRGLGELISMGDSITVTLLAAHQALGKGIILA 156
Db	67	LTTQAGELGLGADVPEVGGGLGDKISSIITKEFSRAGGFSLSYGAHVGGISLPIVF 126
Qy	157	GTEQAKYLPKLASGHIAAFCLTEPASGDAASIRSRATLSEDKKHLYLNGSKVWITN 216
Db	127	GNEDQKKYLPDLATGARIAAYALTPESSGSDALGAKTTAVLNEAGTHYVLNGEKWITN 186
Qy	217	GGLANITVFYAKTEVDSGVDKKTAFIVERDFGVNGKPEDKLGIRGNSCTVEHFE 276
Db	187	SAFADVFIVYAKI-----DG---EHTSAFIVKDFEGVSTGPEKKMGKIGSSTRLLILE 238
Qy	277	NTKIPVENILGEVGDGFKVAMNINLSGRFSVAGLLKRLKLIEMTAETACTKQFNKRL 336
Db	239	DALVPKENLGEVKGKHVIAFINILNVRGYKLGVCIGGSKRGIELAAKYANERKQFKPI 298
Qy	337	SEGLIQEKPALMAQKAYVMESMTYLTAGMLDQ-----PGPPDCSIEA 379
Db	299	AKFTLIQEKLANMAVTTAAESSIYRTGGLFEDRLGGLSEEQOQKDGREVAKAIAEYAE 358
Qy	380	AMVKVFSEAAWQCVSEALQITGLGYTRDYPYERILDRTRILLIPEGTNEILRMVYALT 439
Db	359	SLNKVVGSEALDFVADSAVQIHGGYGFMAEYETMYRDSRINRIPEGTNEILRMVYALT 418
Qy	440	GLOHA--GRI--LTTIRIHELQAKYSTVMDTVGRRLRDSLGRITVDLGLTGNHGVVHPSLAD 496
Db	419	ILRKAMGELPFLEKATALQELMMLMPQEVG-----LONLFS--LSQLDK 588
Qy	497	SANKFEENTYCFGRVETVETLL-----RGKTIMEQLVLRVANILINLYGMTAVLSR 549
Db	451	-DEPLEQEKYLLSMAKVKVFIAGTGAQTYGKELQOQELLANVADISVIFSIVILR 509
Qy	550	ASRST-RIGLRNHDHVELLANTFCVEAY-----LONLFS--LSQLDK 588
Db	510	TEKAIIRNGLKEAKQKALQVFCQAFNRTEAHAKESLVAMQSGDTRMTSILRKLTR 569
Qy	589	YAPENLDPQIKVKSQILEKRAYI 612
Db	570	HTPINVIATKREIAGRILEARYV 593
RESULT 13		
ID	O32176	PRELIMINARY; PRT; 594 AA.
AC	O32176;	
DT	01-JAN-1998 (TremBLrel. 05, Created)	
DT	01-JAN-1998 (TremBLrel. 05, Last sequence update)	
DT	01-MAR-2003 (TremBLrel. 23, Last annotation update)	
DE	YUSJ protein.	
GN	YUSJ.	
OS	Bacillus subtilis.	
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
OX	NCBI_TaxID=1423;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=168;	
RX	MEDLINE=98044033; PubMed=9384377;	
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,	
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,	
RA	Borriss R., Bourschi L., Brans A., Braun M., Brignell S.C., Bron S.,	
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,	
RA	Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,	
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,	
RA	Entian K.D., Etkington J., Fabret C., Ferrari E., Foulger D.,	
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,	
RA	Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,	
RA	Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,	
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,	
RA	Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,	
RA	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,	
RA	Kurita K., Lapidus A., Lardinis S., Lauber J., Lazarevic V.,	
RA	Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,	
RA	Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,	
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,	
RA	Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,	
RA	Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,	
RA	Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,	
RA	Sato T., Scantian E., Schleich S., Schroeter R., Scoffone F.,	
RA	Sekiguchi J., Sekowska A., Seror S.J., Serrit P., Shin B.S., Soldo B.,	
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,	
RA	Takeuchi M., Tamakoshi A., Tanaka T., Terstra P., Tognoni A.,	
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,	
RA	Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegeger T.,	
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,	
RA	Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,	
RT	"The complete genome sequence of the gram-positive bacterium Bacillus subtilis."	
RL	Nature 390:249-256(1997).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=168;	
RA	Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,	
RL	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; Z99120; CAB15271.1; -	
DR	HSSP; P26440; 11VH.	
DR	InterPro; IPRO06089; Acyl-CoA_dh.	
DR	InterPro; IPRO06090; Acyl-CoA_dh_C.	
DR	InterPro; IPRO06091; Acyl-CoA_dh_M.	
DR	InterPro; IPRO06092; Acyl-CoA_dh_N.	
DR	Pfam; PF00441; Acyl-CoA_dh; 1.	
DR	Pfam; PF02770; Acyl-CoA_dh; 1.	
DR	Pfam; PF02771; Acyl-CoA_dh_N; 1.	
DR	PROSITE; PS00072; ACYL_COA_DH_1; 1.	
DR	PROSITE; PS00073; ACYL_COA_DH_2; 1.	
KW	Complete proteome.	
SQ	Complete proteome. 594 AA; 65335 MW; F59BED1AA0D903CE CRC64;	
Query Match		25.9%; Score 815.5; DB 16; Length 594;
Best Local Similarity		33.4%; Pred. No. 5.2e-49;
Matches		201; Conservative 97; Mismatches 202; Indels 101; Gaps 12;
Qy	59	VSODELNEINQFLGPVEKFTTEVDSRKIDQEGKIPDETLKLSLGLFLOVPEYGG 118
Db	45	IEQDVLPHI-----DDIENHOFESVRL-----LKKAGELGLGADVPEYGG 88
Qy	119	GFSTMYSLRGLGELISMGDSITVTLLAAHQALGKGIILAGTTEQAKYLPKLASGEHIAAF 178
Db	89	GLDKISSALITKEFSRAGSFLSYGAHVGGISLPIVFFGSEQKKYLPGLASGEKIAAY 148
Qy	179	CLTEPASGDAASIRSRATLSEDKKHLYLNGSKVWITNGLIANIFVFAKTEVVDSDGV 238
Db	149	ALTEPGSGDALGAKTTAVLNEAGTHYVLNGEKWITNSAFADVFVYAKV-----DG--- 201
Qy	239	KDKITAFIVERDFGVNGKPEDKLGIRGNSCTVEHPTKIPVENILGEVGDGFKVAMN 298
Db	202	-DKFSAFIVKDFEGVSTGPEKKMGKIGSSTRLLDQAEVPEKNLGEIGKHVIAFN 260
Qy	299	ILNSGRFSVAGLLKRLKLIEMTAETACTKQFNKRLSEFGLIOEKFALMAQKAYVMES 358
Db	261	ILNIGRYKLAVGTIGASKRVIELSAAYANORROFKPTIAGLSLTQEKIGTMASRLYAMES 320
Qy	359	MTYLTAGMLDQ-----PGPPDCSIEAAMVKVFSSEAQCVCSEALQIL 401
Db	321	SVYRTVGLFEDNMSQFTAEIDLKGRQRTAKSTAETAEVIECSLNKVFSGSTLIDYDEGVQIH 380
Qy	402	GGLGYTRDYPYERILDRTRILLIPEGTNEILRMVYALTGLOHA--GRI--LTTIRIHELQAK 458
Db	381	GGYGFMAEYEVERAYRDSRINRIPEGTNEILRMVYALTGLOHA--GRI--LTTIRIHELQAK 440

Qy 459 KVTVMDTVGRRLRDSLGRTVDLGLTGNHGVHPSLADSNKFEENTYCFGTRTVETLL- 517
Db :
441 LMMLEPEPG-----SGVL-----EQEYIVKOAKKIALFA 471
Qy 518 -----RFGKTINEBOLVLRKVANILINLYGMTAVLSRASRSIRI--GLRNHDHEVLANT 570
Db : : | | | | | : : : : : | :
472 AGUAAQKYKAIDREOEILNVNADVSNVYAMESAVLRTEKALAQAQKAQKVLYTEI 531
Qy 571 FCVEAYLQ-----NLFSLSQLDKVAPENLDEQTKVKYSQQILEKRA 610
Db :
532 FVOEAFNEIEAHKESLIAMERGDSLUMMLSALRKLTRVTPKNVIOKKREAAGIFEAEK 591
Qy 611 Y 611
Db :
592 Y 592

RESULT 14
Q8ENT4 PRELIMINARY; PRT; 594 AA.

ID	Q8ENT4
AC	Q8ENT4;
DT	01-WAR-2003 (TremBLrel. 23, Created)
DT	01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT	01-WAR-2003 (TremBLrel. 23, Last annotation update)
DE	Acetyl-CoA dehydrogenase.
GN	O2393
OS	Oceanobacillus thelyensis.
OC	Bacteria; Firmicutes; Bacillales; Oceanobacilllus.
OX	NCBI_TaxID=182710;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-HTE831 / DSM 14371 / JCM 11309;
RX	MEDLINE=22220767; PubMed=12235376;
RA	Takami H., Takaki Y., Uchiyama I.;
RT	"Genome sequence of Oceanobacillus thelyensis isolated from the Ihaya Ridge and its unexpected adaptive capabilities to extreme environments.";
RT	Nucleic Acids Res. 30:3927-3935(2002).
RL	EMBL; AP004601; BAC14349.1; ..
KW	Complete proteome.
SQ	SEQUENCE 594 AA; 65915 MW; E0E1F2012C026B92 CRC64;

Query Match
Best Local Similarity 32.3%; Score 789.5; DB 16; Length 594;
Matches 202; Conservative 102; Mismatches 224; Indels 97; Gaps 12;

Qy	47 KIKKEVFPPPEVS-----QDELNEINFLGPVEKFFTEVDSRKID----QECKIPD 95
Db	: : : : : : : : :
Qy	7 KIFGGALVEDITGEDIIPTEDFTDEHKMIAKTTEDFDVLGEV-VPKIDLENHEFEHSV 65
Db	
Qy	96 ETLEKLKSLGFLGLOVPEEVGGFGFNNTMYSLRGELISMDGSTVTLAAHQAIUKGIIL 155
Db	
Qy	66 DLLKKAGDLGLLCADVPQYGGALDKDISSLTITEKFSRAGGSFVTHGAHVIGSLPIVF 125
Db	
Qy	156 ACTEEOKAKYLPKLAGEHHIAFCLETPASGSDAASIRSRTLSEDKHYILNGSKVWIT 215
Db	
Qy	126 FGNDQEKNYLPKLATGETELLAAYALTEPSSGSDALCAKTKLANDAGTHYLNGEKQWIT 185
Db	
Qy	216 NGGLANIYTFVFAKTEVVDDSGSVKDKITAFIVERDFGVTNGKPDKGIRSNCTCVHF 275
Db	
Qy	186 NSAFADVFIYAKI-----DG---EHFTAFIVEREFPGVSTGPEKKMGKIKSSSRTRLVL 237
Db	: : : : : : : : : : : : : : : : : : : : : : : :
Qy	276 ENTKIPVENILGEVGDFKVMANNILNSFRSMGVSVAGLLKRLEMTAEYACTRKQFNKR 335
Db	: : : : : : : : : : : : : : : : : : : : : : : :
Qy	238 EDAAVPIENLLCEKGHGHIYAIFNILNVGYKLAVCGVGSRAIELAAKIYNERNKQFNTP 297
Db	: : : : : : : : : : : : : : : : : : : : : : : :
Qy	336 LSEFGLIQKFALMAQKAYVMESWTYLFAGMLDQ-----PGFPDCSIE 378
Db	: : : : : : : : : : : : : : : : : : : : : : : :
Qy	298 ISSFNLTKEKATMAAGIYANESSYVRTVGLFEQRMGALTDBQLKDGREVARSAEYQIE 357
Db	: : : : : : : : : : : : : : : : : : : : : : : :
Qy	379 AAMVKVFSSEAQCSEALOTGLGLGYTDQYPYERILDRTRILLIFEGTNELLRMVAL 438
Db	: : : : : : : : : : : : : : : : : : : : : : : :

```
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
DR EMBL; AK012773; BAB28460.2; -.
FT NON_TER 1
SQ SEQUENCE 169 AA; 19002 MW; D1B743E8E7605A76 CRC64;

Query Match 23.3%; Score 734; DB 11; Length 169;
Best Local Similarity 88.0%; Pred. No. 4.1e-44;
Matches 146; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 454 ELKQAKYSTVMDTVGRRLRSLGRTVDLGLTGNHGVVHPSLADSANKPFENTYCFGRIVE 513
DB 2 ELKSGNVTVMETIGRKLRSLGRTVDLGLTGVVHPSLGSANKLEENHYFGRTVE 61

QY 514 TLLRFCKTMEEQVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCV 573
DB 62 TLLRFCKNIVEEQVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHEVLANTFCV 121

QY 574 EAYLQNFLSLSQDKYAPENLDEQIKKVSQQLKRAYICAHPLDR 619
DB 122 EAYFNFLSLSQDKNAPENLDEQIKKVSQQLKRAYICAHPLDR 167
```

Search completed: September 3, 2003, 13:30:02
Job time : 113 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:17:21 ; Search time 85 seconds
(without alignments)
1159.638 Million cell updates/sec

Title: US-09-945-326-2

Perfect score: 3153

Sequence: 1 MSGCGLFLRTAAARACRGL.....SQILEKRAYICANPLDRTC 621

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3153	100.0	621	22 AAB73691	Human oxidoreductase
2	3153	100.0	621	22 AAB94839	Human protein sequ
3	3153	100.0	621	23 AAE21680	Human acyl dehydro
4	3153	100.0	628	22 AAU23008	Novel human enzyme
5	2864	90.8	565	21 AAB41800	Human ORFX ORF1564
6	2526	80.1	498	22 AAB94077	Human protein sequ
7	1537	48.7	306	22 AAU23012	Novel human enzyme
8	1360	43.1	655	23 ABB06992	Mouse very long ch
9	1356.5	43.0	653	23 ABB06991	Rat very long ch

10	1345	42.7	655	23 ABB06993	Bovine very long c
11	1308	41.5	655	23 ABB06988	Human MD25 (VICAD)
12	1295.5	41.1	655	22 AAB63264	Drosophila melanog
13	980	31.1	202	22 AAU23009	Novel human enzyme
14	774	24.5	169	24 ABB76260	Human GENSER prote
15	602	19.1	414	22 ABB60259	Drosophila melanog
16	598	19.0	405	22 ABB60897	Drosophila melanog
17	595.5	18.9	409	23 AAE19940	Rice isovaleryl-Co
18	594.5	18.9	415	22 ABB71329	Drosophila melanog
19	588.5	18.7	427	23 AAE19941	Soybean isovaleryl
20	577	18.3	382	21 AAB26154	B. cereus zwitterm
21	573	18.2	639	22 ABB64159	Drosophila melanog
22	573	18.2	639	23 ABB70023	Larval viability a
23	558	17.7	420	22 ABB62543	Drosophila melanog
24	555	17.6	386	22 AAU44325	Propionibacterium
25	549.5	17.4	409	20 AAW97796	Arabidopsis thalia
26	549.5	17.4	409	21 AAG53783	Arabidopsis thalia
27	547.5	17.4	421	23 AAG66900	Human medium chain
28	546.5	17.3	409	21 AAG07514	Arabidopsis thalia
29	531	16.8	107	22 AAU23606	Arabidopsis thalia
30	525.5	16.7	337	21 AAG07515	Novel human enzyme
31	525.5	16.7	337	21 AAG53784	Arabidopsis thalia
32	525	16.7	419	22 ABB61315	Arabidopsis thalia
33	484	15.4	309	21 AAG07516	Drosophila melanog
34	484	15.4	309	21 AAG53785	Arabidopsis thalia
35	480	15.2	513	22 AAE05866	Arabidopsis thalia
36	479	15.2	415	21 AAY92516	Pseudomonas stutze
37	479	15.2	415	22 AAM79247	Human OXRE-13, Ho
38	479	15.2	415	23 ABB61864	Human protein SEQ
39	479	15.2	429	21 AAM80231	prostate cancer-as
40	470	14.9	371	21 AAB45211	Human protein SEQ
41	456	14.5	270	23 AAE19339	Corn isovaleryl-Co
42	452.5	14.4	376	22 ABB52774	Escherichia coli p
43	431.5	13.7	301	21 AAB40259	Human ORFX ORF23 p
44	431.5	13.7	301	23 ABB32054	Human dehydrogenas
45	421.5	13.4	378	21 AAB45210	Gene 38 human secr

ALIGNMENTS

RESULT 1
AAB73691
ID AAB73691 standard; Protein; 621 AA.
XX
AC AAB73691;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human oxidoreductase protein ORP-24.
XX

KW Human oxidoreductase protein; ORP; cell proliferative disorder;
KW arteriosclerosis; cirrhosis; psoriasis; cancer; endocrine disorder;
KW diabetes mellitus; diabetes insipidus; dwarfism; hirsutism; amenorrhea;
KW osteoporosis; metabolic disorder; obesity; phenylketonuria;
KW hypercholesterolemia; reproductive disorder; infertility;
KW ovulatory defect; menstrual cycle defect; endometriosis;
KW polycystic ovary disease; spermatogenesis disruption; impotence;
KW neurological disorder; epilepsy; stroke; Alzheimer's disease;
KW Huntington's disease; Parkinson's disease; Creutzfeldt-Jakob disease;
KW meningitis; cerebral palsy; muscular dystrophy; mood disorder; anxiety;
KW schizophrenic disorder; infection; autoimmune disorder;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS; asthma;
KW allergy; Crohn's disease; atopic dermatitis; gout; multiple sclerosis;
KW rheumatoid arthritis; ulcerative colitis; drug screening;
XX toxicity screening; transgenic animal; SNP detection; gene therapy.

Homo sapiens.

WO200144448-A2.

XX 21-JUN-2001.

of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95993 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 621 AA;

Query Match 100.0%; Score 3153; DB 22; Length 621;
Best Local Similarity 100.0%; Pred. No. 1.8e-274;
Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSCGGLPRTAAARACRGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPPEVS 60
Db 1 MSCGGLPRTAAARACRGLVSTANRRLLRTSPVRAFAKELFLGKIKKKEVFPPEVS 60

QY 61 QDELNEINQFLGPVEKFEFEVDORSKIDQEGKIPDETLEKLSGLGLQVPEYGGGLG 120
Db 61 QDELNEINQFLGPVEKFEFEVDORSKIDQEGKIPDETLEKLSGLGLQVPEYGGGLG 120

QY 121 SNTMSYRLGIIISMDGSIITVLAHQAIIGLKGIIILAGTEQKAKYLPKLAGSHEHIAFCL 180
Db 121 SNTMSYRLGIIISMDGSIITVLAHQAIIGLKGIIILAGTEQKAKYLPKLAGSHEHIAFCL 180

QY 181 TEPASGSDAASIRSRATLSDEKHHYILNGSKVMTNGGLANIETVFAKTEVWDSGVKD 240
Db 181 TEPASGSDAASIRSRATLSDEKHHYILNGSKVMTNGGLANIETVFAKTEVWDSGVKD 240

QY 241 KITAFIVERDFGGVTKGPKDKLGRSNTCEVHFENTKIPVENILGEVGDGKFAVNIL 300
Db 241 KITAFIVERDFGGVTKGPKDKLGRSNTCEVHFENTKIPVENILGEVGDGKFAVNIL 300

QY 301 NSGRFSGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFPALMAQKAYVMSMT 360
Db 301 NSGRFSGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFPALMAQKAYVMSMT 360

QY 361 YLTAGMLDQGFPCDSTEAAWVFSSEAAWQCVSEALQILGLGLGYTRDYPYERILDR 420
Db 361 YLTAGMLDQGFPCDSTEAAWVFSSEAAWQCVSEALQILGLGLGYTRDYPYERILDR 420

QY 421 ILLIFECTNEILRMVIALTGLOHAGRLITRIHELKQAKYSTVMDTVGRRLRDSLGRTVD 480
Db 421 ILLIFECTNEILRMVIALTGLOHAGRLITRIHELKQAKYSTVMDTVGRRLRDSLGRTVD 480

QY 481 LGLTGNGHVHPSLAASANKFENTYCFGRVTETLLRLFGKTIEMESOLVLRVANILINL 540
Db 481 LGLTGNGHVHPSLAASANKFENTYCFGRVTETLLRLFGKTIEMESOLVLRVANILINL 540

QY 541 YGMTAVLSRASRSIRGLRNHDHVELLANTFCVEAYLQNLFSLSOLDKYAPENLDQIKK 600
Db 541 YGMTAVLSRASRSIRGLRNHDHVELLANTFCVEAYLQNLFSLSOLDKYAPENLDQIKK 600

QY 601 VSOQILEKRAYICAHPLDRTC 621
Db 601 VSOQILEKRAYICAHPLDRTC 621

```

RESULT 3

AAE21680

ID AAE21680 standard; Protein; 621 AA.

XX

AC AAE21680;

XX	16-JUL-2002 (first entry)	Human acyl dehydrogenase DHDR-7 (62112 protein).
XX	Human; dehydrogenase-7; DHDR-7; 62112 protein; CNS disorder; epilepsy; Alzheimer's disease; AS; Pick's disease; differentiation disorder; Huntington's disease; autonomic function disorder; hyperthyroidism; depression; schizophrenia; panic migraine; cardiac related disorder; anxiety; obesity; arteriosclerosis; restenosis; Parkinson's disease; angina; hypertension; cardiomyopathy; arrhythmia; muscle weakness; arterial inflammation; cell proliferation disorder; growth disorder; diabetes mellitus; hypertension; migration disorder; gene therapy; fertility disorder; autoimmune disorder; metabolic disorder; ataxia; cancer; hepatic disease; acyl-CoA dehydrogenase; beta-oxidation; enzyme.	
XX	Homo sapiens.	
OS	Key Domain	
XX	Location/Qualifiers	
PH	1..36	
FT	/note= "Dehydrogenase domain"	
FT	23..123	
FT	/note= "Dehydrogenase precursor domain"	
FT	37..426	
FT	/note= "Acyl-CoA very long chain dehydrogenase domain"	
FT	70..432	
FT	/note= "Dehydrogenase-related long acyl-CoA chain oxidoreductase domain"	
FT	74..448	
FT	/note= "Short chain related acyl-CoA dehydrogenase specific domain"	
FT	85..438	
FT	/note= "Acyl CoA dehydrogenase domain"	
FT	85..177	
FT	/note= "Acyl CoA dehydrogenase middle domain"	
FT	147..431	
FT	/note= "ACD-3 acyl-CoA dehydrogenase domain"	
FT	172..349	
FT	/note= "Oxidoreductase acyl-CoA dehydrogenase family domain"	
FT	179..286	
FT	/note= "Acyl CoA dehydrogenase C-terminal domain"	
FT	179..191	
FT	/note= "Acyl CoA dehydrogenase signature 1"	
FT	207..604	
FT	/note= "Acyl-CoA oxidase dehydrogenase oxidoreductase flavoprotein domain"	
FT	290..441	
FT	/note= "Acyl CoA dehydrogenase N-terminal domain"	
FT	399..418	
FT	/note= "Acyl CoA dehydrogenase signature 2"	
FT	408..611	
FT	/note= "Dehydrogenase butyryl domain"	
FT	432..580	
FT	/note= "Polysaccharide deacetylase domain"	
FT	438..621	
FT	/note= "Very long chain dehydrogenase domain"	
FT	460..476	
FT	/note= "Sugar transport protein signature"	
XX	WO200218582-A2.	
XX	07-MAR-2002:	
XX	31-AUG-2001; 2001WO-US27186.	
XX	31-AUG-2000; 2000US-229831P.	
XX	(MILL-) MILLENNIUM PHARM INC.	

XX Meyers R, Hunter JJ;
XX WPI: 2002-329775/36.
DR N-PSDB; AAD34006.
XX
PT New human dehydrogenase polypeptide for diagnosing and treating
PT dehydrogenase-7 associated diseases and disorders e.g. Alzheimer's
PT disease, and to identify modulators of therapeutic use
XX
PS Claim 13; Fig 1; 124pp; English.
XX
CC The invention relates to human dehydrogenase (DHDR)-7 polypeptides
CC referred to as 62112 and nucleic acid molecules encoding such
CC polypeptides. DHDR-7 is an acyl-CoA dehydrogenase, a mitochondrial
CC flavoprotein which catalyses the first step of fatty acid beta-
CC oxidation. Sequences of the invention and their antibodies are
CC useful for treating a disorder, disease or condition which is caused
CC by misregulation (e.g. downregulation or upregulation) of DHDR activity.
CC Examples of disorders include central nervous system (CNS) disorders
CC e.g. Alzheimer's disease (AS), Pick's disease, Parkinson's, Huntington's
CC disease, Gilles de la Tourette's syndrome, multiple sclerosis, epilepsy,
CC amyotrophic lateral sclerosis and Creutzfeldt-Jakob disease; autonomic
CC function disorders e.g. hypertension, depression, schizophrenia, panic
CC migraine, anxiety, obsessive-compulsive disorder and obesity; cardiac
CC related disorder e.g. arteriosclerosis, ischaemia reperfusion injury,
CC restenosis, arterial inflammation, angina, hypertension, cardiomyopathy
CC and arrhythmia; disorders of musculus skeletal system paralysis, muscle
CC weakness e.g. ataxia; cell proliferation, growth, differentiation or
CC migration disorders e.g. cancer, diabetes mellitus, hypothyroidism or
CC hyperthyroidism; reproductive or fertility disorders; autoimmune or
CC or immune deficiency disorders; hepatic disease or dysfunction and
CC metabolic disorders. They are used for screening assays, predictive
CC medicine e.g. diagnostic assays, prognostic assays, monitoring clinical
CC trails, and pharmacogenetics. Polypeptides of the invention are used
CC to identify modulators that modulate their activity. Polynucleotides
CC of the invention are used in gene therapy. The present sequence is
CC human DHDR-7.
XX
XX Sequence 621 AA:
SQ
Query Match 100.0%; Score 3153; DB 23; Length 621;
Best Local Similarity 100.0%; Pred. No. 1.8e-274;
Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSGCGFLRTTAAARACGLVYSTANRLRLTSPVRAFAKELFLGKIKKKEVPFPEVS 60
Db 1 MSGCGFLRTTAAARACGLVYSTANRLRLTSPVRAFAKELFLGKIKKKEVPFPEVS 60
QY 61 QDELNEINQFLGPVEKFTTEEVDSRKIDQEGKIPDETLEKLSGLFLGLOVPEEYGGIGF 120
Db 61 QDELNEINQFLGPVEKFTTEEVDSRKIDQEGKIPDETLEKLSGLFLGLOVPEEYGGIGF 120
QY 121 SNTMTSRGELIISMGSTITVLAHQAGLKGIIILAGTEEOAKYLPKLASGEHIAAFL 180
Db 121 SNTMTSRGELIISMGSTITVLAHQAGLKGIIILAGTEEOAKYLPKLASGEHIAAFL 180
QY 181 TEPASGSDAASIRSRATLSEDKKKHYLLNGSKVWITNGGLANFTVFAKTEVVDSGSKVD 240
Db 181 TEPASGSDAASIRSRATLSEDKKKHYLLNGSKVWITNGGLANFTVFAKTEVVDSGSKVD 240
QY 241 KITAFIVERDFGVNGKPEDKLGIRGNTCEVHENTKIPVENILGEVGDGFKVAMNLT 300
Db 241 KITAFIVERDFGVNGKPEDKLGIRGNTCEVHENTKIPVENILGEVGDGFKVAMNLT 300
QY 301 NSGRFSMGSVVAGLLKRLIEMTAETAEYACTRKQFNKRLSEFGLTOEKFPALMAQKAYVWESMT 360
Db 301 NSGRFSMGSVVAGLLKRLIEMTAETAEYACTRKQFNKRLSEFGLTOEKFPALMAQKAYVWESMT 360
QY 361 YLTAGMLDQPGPDGSIEMAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDR 420
Db 361 YLTAGMLDQPGPDGSIEMAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDR 420

QY 421 ILLIFEGTNEILRMVIALTGLOHAGRIILTRIHHELKQAKVSTVMDTVGRRRLDSIGRTVD 480
Db 421 ILLIFEGTNEILRMVIALTGLOHAGRIILTRIHHELKQAKVSTVMDTVGRRRLDSIGRTVD 480
QY 481 LGLTGNHGVVHPSLADSNKFEENTYCFGRVETILLRFGKTIMEEQLVLRVANILNL 540
Db 481 LGLTGNHGVVHPSLADSNKFEENTYCFGRVETILLRFGKTIMEEQLVLRVANILNL 540
QY 541 YGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQDKYAPENLDEQIKK 600
Db 541 YGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQDKYAPENLDEQIKK 600
QY 601 VSQILEKRAYICAHPLDRTC 621
Db 601 VSQILEKRAYICAHPLDRTC 621
RESULT 4
AAU23008
ID AAU23008 standard; Protein; 628 AA.
XX
AC AAU23008;
XX
DT 17-DEC-2001 (first entry)
DE
XX Novel human enzyme polypeptide #94.
XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotropic; anticoagulant.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01239.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI: 2001-465566/50.

N-PSDB: AAS40878.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases

Claim 11; SEQ ID No 1004; 1180pp; English.

The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders, including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polynucleotides of the invention can also be used in gene therapy. AAU22915-AAU23814 represent the novel human enzyme polypeptides of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Sequence 628 AA;

Query Match 100.0%; Score 3153; DB 22; Length 628;
Best Local Similarity 100.0%; Pred. No. 1.8e-274;
Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Tue Sep 9 10:14:37 2003

us-09-945-326-2.rag

```
QY 1 MSGCGFLRTTAAARACGLVSTANRLLRTSPVRAFAKELFLGKTKKEVPFPPEVS 60
Db 8 MSGCGFLRTTAAARACGLVSTANRLLRTSPVRAFAKELFLGKTKKEVPFPPEVS 67
QY 61 QDELNEINQFLGPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLFLGLQVPEEYGG 120
Db 68 QDELNEINQFLGPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLFLGLQVPEEYGG 127
QY 121 SNTMYSRLEIISMDGSIITVLAHAQAIGLGIILAGTEEOAKYLPKLASGEHIAFCL 180
Db 128 SNTMYSRLEIISMDGSIITVLAHAQAIGLGIILAGTEEOAKYLPKLASGEHIAFCL 187
QY 181 TEPASGSDAASRSRATLSEDKKHVILNGSKVWIITNGGLANFTVFAKTEVVDSGSKVD 240
Db 188 TEPASGSDAASRSRATLSEDKKHVILNGSKVWIITNGGLANFTVFAKTEVVDSGSKVD 247
QY 241 KITAFIVERDFGGVTKGPKEDKLGIRGSGNTCEVHFENTKIPVENILGEVGDGFKVAMN 300
Db 248 KITAFIVERDFGGVTKGPKEDKLGIRGSGNTCEVHFENTKIPVENILGEVGDGFKVAMN 307
QY 301 NSGRFSMGSVAGLLKRLIEMTAETAEYACTRKQFNKRLSEFGLIOEKFALMAQKAYVME 360
Db 308 NSGRFSMGSVAGLLKRLIEMTAETAEYACTRKQFNKRLSEFGLIOEKFALMAQKAYVME 367
QY 361 YLTAGMLDQPGPDGCSIEAAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDR 420
Db 368 YLTAGMLDQPGPDGCSIEAAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDR 427
QY 421 ILLIFGTEINILRMVYALTGLQAGHRLITRHELKQAKVSVMTVGBRLDSLGRVTD 480
Db 428 ILLIFGTEINILRMVYALTGLQAGHRLITRHELKQAKVSVMTVGBRLDSLGRVTD 487
QY 481 LGITGNHGVVHPSLADSANKFENTYCFGRVETLLRFGKTIEMEQVLKRVANILINL 540
Db 488 LGITGNHGVVHPSLADSANKFENTYCFGRVETLLRFGKTIEMEQVLKRVANILINL 547
QY 541 YGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQLDKYPENLDEQIKK 600
Db 548 YGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQLDKYPENLDEQIKK 607
QY 601 VSQIILEKRAYICAHPLDRFC 621
Db 608 VSQIILEKRAYICAHPLDRFC 628
```

RESULT 5

AB41800

ID AB41800 standard; Protein; 565 AA.

XX AC AAB41800;

DT 08-FEB-2001 (first entry)

XX DE Human ORF1564 polypeptide sequence SEQ ID NO:3128.

XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnerable; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antithyroid; antianemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive.

XX OS Homo sapiens.

XX WO200058473-A2.

PN

```
XX 05-OCT-2000.
PD 31-MAR-2000; 2000WO-US086621.
XX PF 31-MAR-1999; 99US-0127607.
XX PR 02-APR-1999; 99US-0127636.
XX PR 05-APR-1999; 99US-0127728.
XX PR 30-MAR-2000; 2000US-0540763.
XX PA (CURA-) CURAGEN CORP.
XX XX Shimkets RA, Leach M;
PI WPI; 2000-602362/57.
XX DR N-PSDB; AAC76009.
XX XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX Claim 11; Page 2345-2346; 5507pp; English.
XX CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX CC
XX SQ Sequence 565 AA;
Query Match 90.8%; Score 2864; DB 21; Length 565;
Best Local Similarity 100.0%; Pred. No. 1.5e-248; Indels 0; Gaps 0;
Matches 564; Conservative 0; Mismatches 0;
QY 58 EVSQDELNEINQFLGPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLFLGLQVPEEYGG 117
Db 2 EVSQDELNEINQFLGPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLFLGLQVPEEYGG 61
QY 118 LGFSNTMYSRLEIISMDGSIITVLAHAQAIGLGIILAGTEEOAKYLPKLASGEHIAA 177
Db 62 LGFSNTMYSRLEIISMDGSIITVLAHAQAIGLGIILAGTEEOAKYLPKLASGEHIAA 121
QY 178 FCLTEPASGSDAASRSRATLSEDKKHVILNGSKVWIITNGGLANFTVFAKTEVVDSG 237
Db 122 FCLTEPASGSDAASRSRATLSEDKKHVILNGSKVWIITNGGLANFTVFAKTEVVDSG 181
QY 238 VKDKITAFIVERDFGGVTKGPKEDKLGIRGSGNTCEVHFENTKIPVENILGEVGDGFKVAM 297
Db 182 VKDKITAFIVERDFGGVTKGPKEDKLGIRGSGNTCEVHFENTKIPVENILGEVGDGFKVAM 241
QY 298 NILNSGRFSMGSVAGLLKRLIEMTAETAEYACTRKQFNKRLSEFGLIOEKFALMAQKAYVME 357
Db 242 NILNSGRFSMGSVAGLLKRLIEMTAETAEYACTRKQFNKRLSEFGLIOEKFALMAQKAYVME 301
QY 358 SMTYLTAGMLDQPGPDGCSIEAAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILR 417
Db
```


Db 302 SMTYLTAGMLDQPGFPCDSIEAAMVKVFSSAAQCVCSEALQILGGLGYTRDYPYRILR 361
 Qy 418 DTRILLIFEGTNEILRMVIALTGLQAGRILTTIRIHELKQAKVSTVMDTVGRRRLDSLGR 477
 Db 362 DTRILLIFEGTNEILRMVIALTGLQAGRILTTIRIHELKQAKVSTVMDTVGRRRLDSLGR 421
 Qy 478 TVDLGLTGNHGVVHPSLADSANKPEENTYCFGRVETLLRFGKTIMEEOLVLRVANIL 537
 Db 422 TVDLGLTGNHGVVHPSLADSANKPEENTYCFGRVETLLRFGKTIMEEOLVLRVANIL 481
 Qy 538 INLYGMPAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSQLDKYAPENLDEQ 597
 Db 482 INLYGMPAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSQLDKYAPENLDEQ 541
 Qy 598 IKVYSQOILEKRAYICAHPLDRTC 621
 Db 542 IKVYSQOILEKRAYICAHPLDRTC 565

RESULT 6

AAB94077 ID AAB94077 standard; Protein: 498 AA.
 XX AC AAB94077;
 XX DT 26-JUN-2001 (first entry)
 XX DE Human protein sequence SEQ ID NO:14271.
 XX KW Human: primer; detection; diagnosis; antisense therapy; gene therapy.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

Claim 8; SEQ ID 14271; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 498 AA;

Query Match 80.1%; Score 2526; DB 22; Length 498;
 Best Local Similarity 100.0%; Pred. No. 3.2e-218;
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 124 MYSRLGEIISMDGSITVTTLAAHQATGLKGIILAGTEEQAKYLPKLASGEHIAAFCLTEP 183
 Db 1 MYSRLGEIISMDGSITVTTLAAHQATGLKGIILAGTEEQAKYLPKLASGEHIAAFCLTEP 60
 Qy 184 ASGSDAASIRSRATLSSEDKKHYILNGSKVWITNGLANIFTVFAKTEVWDSGVKDKIT 243
 Db 61 ASGSDAASIRSRATLSSEDKKHYILNGSKVWITNGLANIFTVFAKTEVWDSGVKDKIT 120
 Qy 244 AFIVERDEGGVTNGKPKDKLGIRSNTECVHFNTKIPVENILGEVGDGFKVAMNINSG 303
 Db 121 AFIVERDEGGVTNGKPKDKLGIRSNTECVHFNTKIPVENILGEVGDGFKVAMNINSG 180
 Qy 304 RFSMGVVAGLLKRLIEMTAETACTRQFNKRLSEFGLIOEKFPALMAQKAYVMSMTYLT 363
 Db 181 RFSMGVVAGLLKRLIEMTAETACTRQFNKRLSEFGLIOEKFPALMAQKAYVMSMTYLT 240
 Qy 364 AGMLDQGFPCDSIEAAMVKVFSSAAQCVCSEALQILGGLGYTRDYPYRILDRTRILL 423
 Db 241 AGMLDQGFPCDSIEAAMVKVFSSAAQCVCSEALQILGGLGYTRDYPYRILDRTRILL 300
 Qy 424 IFEGTNEILRMVIALTGLQAGRILTTIRIHELKQAKVSTVMDTVGRRRLDSLGRTVDLGL 483
 Db 301 IFEGTNEILRMVIALTGLQAGRILTTIRIHELKQAKVSTVMDTVGRRRLDSLGRTVDLGL 360
 Qy 484 TGNHGVVHPSLADSANKPEENTYCFGRVETLLRFGKTIMEEOLVLRVANILINLYGM 543
 Db 361 TGNHGVVHPSLADSANKPEENTYCFGRVETLLRFGKTIMEEOLVLRVANILINLYGM 420
 Qy 544 TAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSQLDKYAPENLDRQIKKVSQ 603
 Db 421 TAVLSRASRSIRIGLRNHDHVEVLANTFCVEAYLQNLFSQLDKYAPENLDRQIKKVSQ 480
 Qy 604 QILEKRAYICAHPLDRTC 621
 Db 481 QILEKRAYICAHPLDRTC 498

RESULT 7

AAU23012 ID AAU23012 standard; Protein: 306 AA.
 XX AC AAU23012;
 XX DT 18-DEC-2001 (first entry)
 XX DE Novel human enzyme polypeptide #98.

KW Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW nephrotropic; anticoagulant.

OS Homo sapiens.

XX XX

PN WO200155301-A2.

XX

Tue Sep 9 10:14:37 2003

us-09-945-326-2.rag

PD	02-AUG-2001.	PR	29-SEP-2000; 2000US-0236368.
XX		PR	29-SEP-2000; 2000US-0236369.
PF	17-JAN-2001; 2001WO-US01239.	PR	29-SEP-2000; 2000US-0236370.
PF		PR	02-OCT-2000; 2000US-0236802.
XX		PR	02-OCT-2000; 2000US-0237037.
PR	31-JAN-2000; 2000US-0179065.	PR	02-OCT-2000; 2000US-0237038.
PR	04-FEB-2000; 2000US-0180628.	PR	02-OCT-2000; 2000US-0237039.
PR	24-FEB-2000; 2000US-0184664.	PR	02-OCT-2000; 2000US-0237040.
PR	02-MAR-2000; 2000US-0186350.	PR	13-OCT-2000; 2000US-0239935.
PR	16-MAR-2000; 2000US-0189874.	PR	13-OCT-2000; 2000US-0239937.
PR	17-MAR-2000; 2000US-0190076.	PR	20-OCT-2000; 2000US-0240960.
PR	18-APR-2000; 2000US-0198123.	PR	20-OCT-2000; 2000US-0241221.
PR	19-MAY-2000; 2000US-0205515.	PR	20-OCT-2000; 2000US-0241785.
PR	07-JUN-2000; 2000US-0209467.	PR	20-OCT-2000; 2000US-0241786.
PR	28-JUN-2000; 2000US-0214886.	PR	20-OCT-2000; 2000US-0241787.
PR	30-JUN-2000; 2000US-0215135.	PR	20-OCT-2000; 2000US-0241808.
PR	07-JUL-2000; 2000US-0216647.	PR	20-OCT-2000; 2000US-0241809.
PR	07-JUL-2000; 2000US-0216880.	PR	20-OCT-2000; 2000US-0241826.
PR	11-JUL-2000; 2000US-0217487.	PR	01-NOV-2000; 2000US-0244617.
PR	11-JUL-2000; 2000US-0217496.	PR	08-NOV-2000; 2000US-0246474.
PR	14-JUL-2000; 2000US-0218290.	PR	08-NOV-2000; 2000US-0246475.
PR	26-JUL-2000; 2000US-0220963.	PR	08-NOV-2000; 2000US-0246476.
PR	26-JUL-2000; 2000US-0220964.	PR	08-NOV-2000; 2000US-0246477.
PR	14-AUG-2000; 2000US-0224518.	PR	08-NOV-2000; 2000US-0246478.
PR	14-AUG-2000; 2000US-0224519.	PR	08-NOV-2000; 2000US-0246523.
PR	14-AUG-2000; 2000US-0225213.	PR	08-NOV-2000; 2000US-0246524.
PR	14-AUG-2000; 2000US-0225214.	PR	08-NOV-2000; 2000US-0246525.
PR	14-AUG-2000; 2000US-0225266.	PR	08-NOV-2000; 2000US-0246526.
PR	14-AUG-2000; 2000US-0225267.	PR	08-NOV-2000; 2000US-0246527.
PR	14-AUG-2000; 2000US-0225268.	PR	08-NOV-2000; 2000US-0246528.
PR	14-AUG-2000; 2000US-0225270.	PR	08-NOV-2000; 2000US-0246532.
PR	14-AUG-2000; 2000US-0225447.	PR	08-NOV-2000; 2000US-0246609.
PR	14-AUG-2000; 2000US-0225757.	PR	08-NOV-2000; 2000US-0246610.
PR	14-AUG-2000; 2000US-0225758.	PR	08-NOV-2000; 2000US-0246611.
PR	14-AUG-2000; 2000US-0225759.	PR	08-NOV-2000; 2000US-0246613.
PR	18-AUG-2000; 2000US-0226279.	PR	17-NOV-2000; 2000US-0249207.
PR	22-AUG-2000; 2000US-0226681.	PR	17-NOV-2000; 2000US-0249208.
PR	22-AUG-2000; 2000US-0226686.	PR	17-NOV-2000; 2000US-0249209.
PR	22-AUG-2000; 2000US-0227182.	PR	17-NOV-2000; 2000US-0249210.
PR	23-AUG-2000; 2000US-0227009.	PR	17-NOV-2000; 2000US-0249211.
PR	30-AUG-2000; 2000US-0228924.	PR	17-NOV-2000; 2000US-0249212.
PR	01-SEP-2000; 2000US-0229287.	PR	17-NOV-2000; 2000US-0249213.
PR	01-SEP-2000; 2000US-0229343.	PR	17-NOV-2000; 2000US-0249214.
PR	01-SEP-2000; 2000US-0229344.	PR	17-NOV-2000; 2000US-0249215.
PR	01-SEP-2000; 2000US-0229345.	PR	17-NOV-2000; 2000US-0249216.
PR	05-SEP-2000; 2000US-0229509.	PR	17-NOV-2000; 2000US-0249217.
PR	05-SEP-2000; 2000US-0229513.	PR	17-NOV-2000; 2000US-0249218.
PR	06-SEP-2000; 2000US-0230437.	PR	17-NOV-2000; 2000US-0249244.
PR	06-SEP-2000; 2000US-0230438.	PR	17-NOV-2000; 2000US-0249245.
PR	08-SEP-2000; 2000US-0231242.	PR	17-NOV-2000; 2000US-0249264.
PR	08-SEP-2000; 2000US-0231243.	PR	17-NOV-2000; 2000US-0249265.
PR	08-SEP-2000; 2000US-0231244.	PR	17-NOV-2000; 2000US-0249297.
PR	08-SEP-2000; 2000US-0231413.	PR	17-NOV-2000; 2000US-0249299.
PR	08-SEP-2000; 2000US-0231414.	PR	17-NOV-2000; 2000US-0249300.
PR	08-SEP-2000; 2000US-0232080.	PR	01-DEC-2000; 2000US-0250160.
PR	08-SEP-2000; 2000US-0232081.	PR	01-DEC-2000; 2000US-0250391.
PR	12-SEP-2000; 2000US-0231968.	PR	05-DEC-2000; 2000US-0251030.
PR	14-SEP-2000; 2000US-0232397.	PR	05-DEC-2000; 2000US-0251988.
PR	14-SEP-2000; 2000US-0232398.	PR	05-DEC-2000; 2000US-0256719.
PR	14-SEP-2000; 2000US-0232399.	PR	06-DEC-2000; 2000US-0251479.
PR	14-SEP-2000; 2000US-0232400.	PR	08-DEC-2000; 2000US-0251856.
PR	14-SEP-2000; 2000US-0232401.	PR	08-DEC-2000; 2000US-0251868.
PR	14-SEP-2000; 2000US-0233063.	PR	08-DEC-2000; 2000US-0251869.
PR	14-SEP-2000; 2000US-0233064.	PR	08-DEC-2000; 2000US-0251989.
PR	14-SEP-2000; 2000US-0233065.	PR	11-DEC-2000; 2000US-0251990.
PR	21-SEP-2000; 2000US-0234223.	PR	05-JAN-2001; 2001US-0259678.
PR	21-SEP-2000; 2000US-0234274.	XX	
PR	25-SEP-2000; 2000US-0234997.	XX	(HUMA-) HUMAN GENOME SCI INC.
PR	25-SEP-2000; 2000US-0234998.	XX	Rosen CA, Barash SC, Ruben SM;
PR	26-SEP-2000; 2000US-0235484.	XX	WPI; 2001-465566/50.
PR	27-SEP-2000; 2000US-0235834.	XX	
PR	27-SEP-2000; 2000US-0235836.	XX	
PR	29-SEP-2000; 2000US-0236327.	XX	
PR	29-SEP-2000; 2000US-0236367.	XX	

QY 394 VSEALQILGGLGYTRDYPYRILDRTRILLIFEGTNEILRMVIALTGLQAHAGRIITRIH 453
 Db 430 TDECIOIMGGFMKPGVERVLRDLRIFRIFEGTNDILRLFLVALQCMQDKGKLSGLGS 489
 QY 454 ELKQ--AKVSTVMDTVGRRLDSLGRVLDGLTGNHGVVHPSLADSANKFPENTYCFGR 511
 Db 490 ALKNPFGNAGLLGAGKQLRRRAGLGLSLS---GLVHPELSRSGELAVRALEQFATV 546
 QY 512 VETLLRFGKTIMEQLVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHEVLLANTF 571
 Db 547 VEAKLIKHKKGIVNEQFLRLADCAIDLYAMVVVLSRASRSLSGHPPTAQHEKMLCDTW 606
 QY 572 CVE--AYLONLESQDKYAPENLDEQIKKVSQOILEKRAYICAHPL 617
 Db 607 CIEAARIREGMAALQSDPWOQE-LYRNFKSISKALVERGGVVTISNPL 653

RESULT 12

ABB63264
 ID ABB63264 standard; Protein; 655 AA.

XX ABB63264;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 16584.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX W0200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL07367.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 16584; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 655 AA;

XX Query Match 41.1%; Score 1295.5; DB 22; Length 655;

XX Best Local Similarity 45.1%; Pred. No. 2.5e-107; Gaps 9;

XX Matches 280; Conservative 104; Mismatches 212; Indels 25; Gaps 9;

QY 14 ARACRGLVV-----STANRRLLRTPPVRAFAKELFLGKIKKKEVFFPEVSQDELNEI 67

Db 44 ASLCROIAHSPKLGAESENRSKEKASEN-ESEMANFRGLSVSSQVFPDVLTAQOKEL 102
 QY 68 -NDFLGPVEKFFTEEDSRKIDQEGKIPDETLEKLSKSLGLFGLQVPEEYGGGFSNTMYS 126
 Db 103 TNSLIDFFERFFSDVNDAAANDANSKIDDTTSTALWELGAFGLQVPSERGGGLNNTQYG 162
 QY 127 RLGEITISM-DGSTVTTLAAHQAIGLGIILAGTTEQAKYLPKLASGEHIAAFCLTEPAS 185
 Db 163 RLCAIVGNDLGLGITITGAHOSIGFGKILLYGTPQEKYLPKVAAEQVAAAFALTEPSS 222
 QY 186 GSDAASRSTRATLSDEKKHYILNGSKVWITNGLANLFTVFAKTEVVD-SDGSVKDKITA 244
 Db 223 GSDAGSIRCAVKSADGKHVYLVNGSKTWISNGGTAETMTVFAQTEQVDPKTKGKDKVTA 282
 QY 245 FIVERDFGGVTKNGPKDLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMNINLSGR 304
 Db 283 FIVERSFGVTKNGPPEKKMGKIKASNTAEVYFEDVKIPINVLGKEGDGFVAMNINLSGR 342
 QY 305 FMSGVAVAGLLKRLIEMTAETAYACTRKQFNKELSEFGLIOEKFALMAQKAYVMESMTYLA 364
 Db 343 FPMGATLSGTMKKICIEQATEHANNRVQFGOKLKNYSIQEKLAQNMNLTQATSMAPTIS 402
 QY 365 GMLDQPGFPCSIIEAAMVKVFSSEAAWQCVSEALQILGGLGYTRDYPYRILDRTRILLI 424
 Db 403 QNMD-AGSKDYLHLEAAISKIYASASAWYVCDEAIQILGGMGYMVDNGLERVLRLDFRI 461
 QY 425 FEGTNEILRMVIALTGLQAHAGRIITRIHELKQ-----AKVSTVMDTVGRRLDSLGR 478
 Db 462 FEGTNDILRLFTALTGTQYAG---SHLKEQLRAFKNPSANGLIFKEASHR-----AAS 513
 QY 479 VDLGLTGNHGVVHPSLADSANKFPENTYCFGRVETILLRFGKTIMEEQLVKRVANILI 538
 Db 514 VDLGGLTDLGSHVVGELLPYAKKTAHCIDLFGQSVEELLRYNKNIVNEQIILLTLANAAI 573
 QY 539 NLYGMTAVLSRASRSIRIGLRNHDHEVLLANTFCVEAYLQNLFSLSQDKYAPENLDEQI 598
 Db 574 DIYAMVVTQSRSSRAVNLNLTQAHELNNKALTATQASDRVIKNLQAATSSHHRSLNEKI 633
 QY 599 KKVSOILEKRAYICAHPLDR 619
 Db 634 STIAKTTLENGGVTTTGILDQ 654

RESULT 13

AAU23009
 ID AAU23009 standard; Protein; 202 AA.

XX AAU23009;

XX 17-DEC-2001 (first entry)

DE Novel human enzyme polypeptide #95.

XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW nephrotropic; anticoagulant.

XX Homo sapiens.

XX W0200155301-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01239.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184564.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217496.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465566/50.

XX N-PSDB; AAS40879.

PT Novel polypeptides and polynucleotides useful for diagnosing,
preventing, treating neural, immune system, muscular, reproductive,
pulmonary, cardiovascular, renal, proliferative disorders and cancerous
diseases

PS Claim 11; SEQ ID No 1005; 1180pp; English.

DR XX The present invention relates to the isolation of novel human enzyme
 CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
 CC encoding them. The enzyme polypeptides of the invention may comprise the
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
 CC isomerases or ligases. The sequences of the invention are useful in the
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of
 CC disorders including hyperproliferative disorders (e.g. cancer),
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
 CC blood-related disorders (e.g. haemophilia), reproductive disorders
 CC (e.g. infertility) and infectious disorders (e.g. influenza). The
 CC polynucleotides of the invention can also be used in gene therapy.
 CC AAU22915-AAU23814 represent the novel human enzyme polypeptides of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 202 AA;
 Query Match 31.1%; Score 980; DB 22; Length 202;
 Best Local Similarity 100.0%; Pred. No. 9.7e-80;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGCGLFLRTTAAARACRGLVSTANRLRLTSPVRAFAKELFLGKIKKEVFPFPEVS 60
 DB 8 MSGCGLFLRTTAAARACRGLVSTANRLRLTSPVRAFAKELFLGKIKKEVFPFPEVS 67
 QY 61 QDELNEINQFLGVPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGIGF 120
 DB 68 QDELNEINQFLGVPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGIGF 127
 QY 121 SNTMYSRLGEIISMDGSIITVLAHQAIIGLKGIIAGTTEOKAKYLPKL 180
 DB 128 SNTMYSRLGEIISMDGSIITVLAHQAIIGLKGIIAGTTEOKAKYLPKL 187
 QY 181 TEPASGSDAASIRSR 195
 DB 188 TEPASGSDAASIRSR 202

RESULT 14
 ABP76260 standard; Protein: 169 AA.
 XX AC ABP76260;
 XX 21-FEB-2003 (first entry)
 XX Human GENSET protein SEQ ID 810.
 XX Cytostatic; antiinflammatory; nootropic; neuroprotective; cardiant;
 KW gastrointestinal; gene therapy; GENSET; heavy metal toxicity; cancer;
 KW inflammatory disease; immune disorder; neuromuscular; toxicity;
 KW central nervous system; cardiovascular; gastrointestinal.
 XX Homo sapiens.
 XX WO200283898-A1.
 XX 24-OCT-2002.
 XX 18-APR-2001; 2001WO-IB00914.
 XX 18-APR-2001; 2001WO-IB00914.
 XX (GEST) GENSET.
 XX Bejanin S, Tanaka H, Dumas Milne Edwards J, Jobert S, Giordano J;
 XX

DR XX WPI: 2003-075548/07.
 XX New GENSET polynucleotides and polypeptides, useful for treating heavy
 PT metal toxicity, cancer, inflammatory diseases, immune disorders, and
 PT the neuromuscular, CNS, cardiovascular or gastrointestinal effects of
 PT the toxicity
 XX Claim 14; Page 690; 735pp; English.
 XX The present invention relates to novel GENSET polynucleotides
 CC (ABZ36404-ABZ36911) encoding polypeptides (ABP75963-ABP76368). The
 CC polynucleotides and polypeptides are useful in screening and diagnostic
 CC assays for abnormal GENSET expression and/or biological activity. They
 CC are also useful for screening of compounds for treating or preventing
 CC GENSET-related disorders, such as heavy metal toxicity, cancer, central
 CC inflammatory diseases, immune disorders, and the neuromuscular, central
 CC nervous system (CNS), cardiovascular or gastrointestinal effects of the
 CC toxicity.
 XX Sequence 169 AA;
 SQ Query Match 24.5%; Score 774; DB 24; Length 169;
 Best Local Similarity 91.7%; Pred. No. 2.5e-61;
 Matches 155; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 MSGCGLFLRTTAAARACRGLVSTANRLRLTSPVRAFAKELFLGKIKKEVFPFPEVS 60
 DB 1 MSGCGLFLRTTAAARACRGLVSTANRLRLTSPVRAFAKELFLGKIKKEVFPFPEVS 60
 QY 61 QDELNEINQFLGVPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGIGF 120
 DB 61 QDELNEINQFLGVPVEKFFTEEVDSRKIDQEGKIPDETLEKLSGLGLQVPEEYGGIGF 120
 QY 121 SNTMYSRLGEIISMDGSIITVLAHQAIIGLKGIIAGTTEOKAKYLPKL 169
 DB 121 SNTMYSRLGEIISMDGSIITVLAHQAIIGLKGIIAGTTEOKAKYLPKL 169

RESULT 15
 ABP60259 standard; Protein: 414 AA.
 XX AC ABP60259;
 XX 26-MAR-2002 (first entry)
 XX Drosophila melanogaster polypeptide SEQ ID NO 7569.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2001; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI: 2001-656860/75.
 DR N-PSDB; ABL04362.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX

Search completed: September 3, 2003, 13:27:31
Job time : 87 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 10:14:14 ; Search time 513.782 seconds
(without alignments)
9788.298 Million cell updates/sec

Title: US-09-945-326-3
Perfect score: 1863
Sequence: 1 atgagcggctgcgggtctt.....acccctggacagacatgc 1863

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25:	/SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1863	100.0	2452	24	AAD34006	Human acyl dehydro
2	1863	100.0	2522	23	ABV21123	Human prostate exp
3	1863	100.0	2522	23	ABV23498	Human prostate exp
4	1863	100.0	2522	23	ABV26965	Human prostate exp
5	1863	100.0	2522	23	ABV29355	Human prostate exp
6	1862.6	100.0	2492	22	AA40878	cDNA encoding nove
7	1861.4	99.9	2440	22	AAH24246	Human oxidoreducta
8	1859.8	99.8	2399	22	AAH16781	Human cDNA sequenc

9	1590.8	90.8	2218	21	AAC76009	Human OREF ORF1564
10	1671.4	89.7	2400	22	AAH15804	Human CDNA sequenc
11	1912.6	49.0	1118	22	AA40882	cDNA encoding nove
12	704.6	37.8	771	22	AAH08631	Human CDNA clone (
13	585	31.4	651	22	AA40879	cDNA encoding nove
14	499.6	26.8	772	20	AA408901	Human validated ca
15	480.4	26.3	835	20	AA215890	Human gene express
16	484.4	26.0	543	25	AB236713	Human GENSET codin
17	460.4	24.7	507	25	AB236843	Human GENSET codin
18	418.4	22.5	431	22	AAF67278	Novel human polynu
19	393.2	21.1	796	22	AAH05329	Human CDNA clone (
20	391.6	21.0	1965	24	ABL50829	Human MD25 (VLCAD)
21	391.6	21.0	2148	24	ABL50834	Human MD25 (VLCAD)
22	388.6	20.9	2117	24	ABK63677	Rat sequence diffe
23	380.2	20.4	434	22	AAF65775	Novel human polynu
24	372.8	20.0	2045	23	ABL12894	Drosophila melanog
25	372.8	20.0	4045	23	ABL12894	Drosophila melanog
26	372.6	20.0	412	22	AAF65875	Novel human polynu
27	358.4	19.2	2186	21	AAC93369	Human secreted pro
28	318.8	17.1	321	22	AA41476	cDNA encoding nove
29	300	16.1	300	20	AA213131	Human gene express
30	283.2	15.2	300	20	AA408613	Human cancer cell
31	259.2	13.9	472	23	ABV44576	Human prostate exp
32	233.8	12.5	252	23	ABV14697	Human prostate exp
33	225.4	12.1	275	23	ABV35775	Human prostate exp
34	221.8	11.9	1024	25	AB283763	Toxicologically re
35	214.6	11.5	312	24	ABK35342	Human cDNA encodin
36	197.8	10.6	462	23	ABV32112	Human prostate exp
37	196.4	10.5	458	23	ABV41049	Human prostate exp
38	191.8	10.3	221	23	ABV01794	Human prostate exp
39	183.4	9.8	185	23	ABV10963	Human prostate exp
40	159.4	8.6	183	24	ABL80414	Human ovarian canc
41	146.6	7.9	241	20	AA439501	Human secreted pro
42	146.6	7.9	1829	24	ABN95811	Gene #2309 used to
43	143.4	7.7	4403765	22	AA199683	Mycobacterium tube
44	143.4	7.7	4411529	22	AA199682	Mycobacterium tube
45	140.4	7.5	1569	23	ABL06885	Drosophila melanog

ALIGNMENTS

RESULT 1
AAD34006
ID AAD34006 standard; CDNA; 2452 BP.

XX AAD34006;

XX AAD34006;

DT 16-JUL-2002 (first entry)

XX Human acyl dehydrogenase DHDR-7 (62112 protein)-encoding CDNA.

XX Human: dehydrogenase-7; DHDR-7; 62112 protein; CNS disorder;
XX epilepsy; Alzheimer's disease; AS; Pick's disease;
XX differentiation disorder; Huntington's disease; autonomic
XX function disorder; hyperthyroidism; depression; schizophrenia;
XX panic migraine; cardiac related disorder; anxiety; obesity;
XX arteriosclerosis; restenosis; Parkinson's disease; angina;
XX hypertension; cardiomyopathy; arrhythmia; muscle weakness;
XX arterial inflammation; cell proliferation disorder; growth disorder;
XX diabetes mellitus; hypertension; migration disorder; gene therapy;
XX fertility disorder; autoimmune disorder; metabolic disorder; ataxia;
XX cancer; hepatic disease; acyl-CoA dehydrogenase; beta-oxidation;
XX gene; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

XX 5'UTR 1..66

FT /*tag= a

FT CDS 67..1932

FT /*tag= b

FT /product= "Human DHDR-7 protein"

FT	/note= "This region is specifically claimed in
FT	claim 1 as SEQ ID NO:3"
FT	3'UTR 1933..2452
FT	/*tag= C
FT	
PN	WO200218582-A2.
PD	07-MAR-2002.
XX	31-AUG-2001; 2001WO-US27186.
XX	31-AUG-2000; 2000US-229831P.
PR	(MILL-) MILLENNIUM PHARM INC.
PA	Meyers R, Hunter JJ;
PI	WPT; 2002-329775/36.
DR	P-PSDB; AAE21680.
XX	
PT	New human dehydrogenase polypeptide for diagnosing and treating
PT	dehydrogenase-7 associated diseases and disorders e.g. Alzheimer's
PT	disease, and to identify modulators of therapeutic use -
XX	
PS	Claim 1; Fig 1; 124pp; English.
CC	The invention relates to human dehydrogenase (DHDR)-7 polypeptides
CC	referred to as G212 and nucleic acid molecules encoding such
CC	polypeptides. DHDR-7 is an acyl-CoA dehydrogenase, a mitochondrial
CC	flavoprotein which catalyses the first step of fatty acid beta-
CC	oxidation. Sequences of the invention and their antibodies are
CC	useful for treating a disorder, disease or condition which is caused
CC	by misregulation (e.g. downregulation or upregulation) of DHDR activity.
CC	Examples of disorders include central nervous system (CNS) disorders
CC	e.g. Alzheimer's disease (AS), Pick's disease, Parkinson's, Huntington's
CC	disease, Gilles de la Tourette's syndrome, multiple sclerosis, epilepsy,
CC	amyotrophic lateral sclerosis and Creutzfeldt-Jakob disease; autonomic
CC	function disorders e.g. hypertension, depression, schizophrenia, panic
CC	migraine, anxiety, obsessive-compulsive disorder and obesity; cardiac
CC	related disorder e.g. arteriosclerosis, ischaemia reperfusion injury,
CC	restenosis, arterial inflammation, angina, hypertension, cardiomyopathy
CC	and arrhythmia; disorders of muscular skeletal system paralysis, muscle
CC	weakness e.g. ataxia; cell proliferation, growth, differentiation or
CC	migration disorders e.g. cancer, diabetes mellitus, hypothyroidism or
CC	hyperthyroidism; reproductive or fertility disorders; autoimmune or
CC	immune deficiency disorders; hepatic disease or dysfunction and
CC	metabolic disorders. They are used for screening assays, predictive
CC	medicine e.g. diagnostic assays, prognostic assays, monitoring clinical
CC	trials, and pharmacogenetics. Polypeptides of the invention are used
CC	to identify modulators that modulate their activity. Polynucleotides
CC	of the invention are used in gene therapy. The present sequence is
CC	cDNA encoding human DHDR-7:
XX	
SQ	Sequence 2452 BP; 616 A; 616 C; 685 G; 535 T; 0 other;
	Query Match 100.0%; Score 1863; DB 24; Length 2452;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGAGCGGTCTGGGGCTCTTCTTGCGCACACCGGCTGCGGCTCGCTGCCTGCGGGGTCTG 60
Dd	67 ATGAGCGGTCTGGGGCTCTTCTTGCGCACACCGGCTGCGGCTCGCTGCCTGCGGGGTCTG 126
QY	61 GTGGTCTCTACCGCAACCGGGGTACTTGGCGCACCGCGGCTGTACGAGCTTCGCC 120
Dd	127 GTGGTCTCTACCGCAACCGGGGTACTTGGCGCACCGCGGCTGTACGAGCTTCGCC 186
QY	121 AAAGAGCTTTCTTAGGCAAAATCAAGAAGAAGATTTCCTCCATTTCCAGAAGTTAGC 180
Dd	187 AAAGAGCTTTCTTAGGCAAAATCAAGAAGAAGATTTCCTCCATTTCCAGAAGTTAGC 246
QY	181 CAAGATGAACCTTAATGAATCAATCAGTTCTTTGGACCCCGTGGAAAAATTCCTACTCAA 240

QY 1321 CTGCAGCATGCCGCGCATCTCTGACTACAGGATCCATGAGCTTAACAGGCCAAAGTG 1380
DB 1387 CTGCAGCATGCCGCGCATCTCTGACTACAGGATCCATGAGCTTAACAGGCCAAAGTG 1446
QY 1381 AGCACATCATGATACCTGTTGGCGGAGGCTTTGGGACTCCCTGGGCGGAACTGTGGAC 1440
DB 1447 AGCACATCATGATACCTGTTGGCGGAGGCTTTGGGACTCCCTGGGCGGAACTGTGGAC 1506
QY 1441 CTGGGCTGACAGCAACCATGAGTTGTGCACCCAGCTTTCGGGACAGTGCCACAAG 1500
DB 1507 CTGGGCTGACAGCAACCATGAGTTGTGCACCCAGCTTTCGGGACAGTGCCACAAG 1566
QY 1501 TTTGAGGAGAACACCTACTCTTTCGGCGGACCGTGGAGACTGCTGCTCCGCTTTGGC 1560
DB 1567 TTTGAGGAGAACACCTACTCTTTCGGCGGACCGTGGAGACTGCTGCTCCGCTTTGGC 1626
QY 1561 AAGACCATCATGAGGAGCAGCTGTTGTAAGCGGGTGGCCAACTCTCATCAACCTG 1620
DB 1627 AAGACCATCATGAGGAGCAGCTGTTGTAAGCGGGTGGCCAACTCTCATCAACCTG 1686
QY 1621 TATGGCATGACGCGCTGCTGTCGGCGGACCGCTCCATCGCATTTGGGCTTCCGCAAC 1680
DB 1687 TATGGCATGACGCGCTGCTGTCGGCGGACCGCTCCATCGCATTTGGGCTTCCGCAAC 1746
QY 1681 CAGGACACAGGTTCTTTCGGCCAAACACTTCTTGGTGAAGCTTACTTGCAGATCTC 1740
DB 1747 CAGGACACAGGTTCTTTCGGCCAAACACTTCTTGGTGAAGCTTACTTGCAGATCTC 1806
QY 1741 TTCAGCTCTCTCAGCTGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAA 1800
DB 1807 TTCAGCTCTCTCAGCTGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAA 1866
QY 1801 GTGTCCACAGATCCTTGAGAGCGAGCTATATCTGTGCCACCTCTGACAGGACA 1860
DB 1867 GTGTCCACAGATCCTTGAGAGCGAGCTATATCTGTGCCACCTCTGACAGGACA 1926
QY 1861 TGC 1863
DB 1927 TGC 1929

RESULT 2
ID ABV21123 standard; cDNA; 2522 BP.
XX
AC ABV21123;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21114.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;

XX

DR WPI: 2001-662795/76.

XX

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

PT

Claim 1; Page 3496; 11750pp; English.

XX

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:

CC

(a) assessing whether a patient is afflicted with prostate cancer;

CC

(b) monitoring the progression of prostate cancer in a patient;

CC

(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

CC

(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC

(e) selecting a composition for inhibiting prostate cancer in a patient;

CC

(f) assessing the prostate cell carcinogenic potential of a compound;

CC

(g) determining whether prostate cancer has metastasized in a patient;

CC

(h) assessing the aggressiveness or indolence of prostate cancer in a patient;

CC

(I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX

Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;

SQ

Query Match

Best Local Similarity 100.0%; Score 1863; DB 23; Length 2522;

DB

Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 ATGAGCGCTCGGGCTCTTCTCGGACACAGCGCTGGGCTCGTGGCTCGCGGGTCTG 60

DB

117 ATGAGCGCTCGGGCTCTTCTCGGACACAGCGCTGGGCTCGTGGCTCGCGGGTCTG 176

QY

61 GTGTCTCTACCGGCAACCGCGGCTACTGCGCACACAGCGCTGTACGAGCTTCGCC 120

DB

177 GTGTCTCTACCGGCAACCGCGGCTACTGCGCACACAGCGCTGTACGAGCTTCGCC 236

QY

121 AAGAGCTTTTCCTAGGCAAAATCAAGAAGAAAGTTTTCCTATTTCCAGAGTTAGC 180

DB

237 AAGAGCTTTTCCTAGGCAAAATCAAGAAGAAAGTTTTCCTATTTCCAGAGTTAGC 296

QY

181 CAAGATGAACCTTAATGAATCAATCAGTCTCTTGGGACCCGTTGGAATAATTTCCAC 240

DB

297 CAAGATGAACCTTAATGAATCAATCAGTCTCTTGGGACCCGTTGGAATAATTTCCAC 356

QY

241 GAGTGGACTCCCGAAAAATTTGACAGGAGGAAAAATCCAGATGAATAATTTGGAGAA 300

DB

357 GAGTGGACTCCCGAAAAATTTGACAGGAGGAAAAATCCAGATGAATAATTTGGAGAA 416

QY

301 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTTGGCTTC 360

DB

417 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTTGGCTTC 476

QY

361 TCCAACACCTACTACTCAAGACTAGGGAGATCATCAGATGGTGGTCCATCCTGTTG 420

DB

477 TCCAACACCTACTACTCAAGACTAGGGAGATCATCAGATGGTGGTCCATCCTGTTG 536

QY

421 ACCCTGGCAGCGCACAGGCTATTGGCTTCAAGGGATCATTTGGCTGGCTGAGGAG 480

DB

537 ACCCTGGCAGCGCACAGGCTATTGGCTTCAAGGGATCATTTGGCTGGCTGAGGAG 596

QY

481 CAGAAAGCCCAATACTTGCCTAACTTGGCTGCGGGAGCACATTTCCAGCTTCTGCC 540

DB

597 CAGAAAGCCCAATACTTGCCTAACTTGGCTGCGGGAGCACATTTCCAGCTTCTGCC 656

QY

541 ACGGAGCCAGCTAGTGGGAGGATGAGGCTTCAATCCGGAGCAGACCACTAAGTGAA 600

DB

657 ACGGAGCCAGCTAGTGGGAGGATGAGGCTTCAATCCGGAGCAGACCACTAAGTGAA 716

QY

601 GACAAGACCTACATCTCAATGCTCAAGTCTGATTTACTAATGAGGACTGGCC 660

PI

Schlegel R, Endege WO, Monahan JE;

Db 717 GACAGAAGCACTACATCTCAATGGCTCCAAAGGCTCTGGATTACTAATGGAGGACTGGCC 776
Qy 661 AATATTTTACTGTGTTTTCGAAGACTGAGTCTGTTGATCTCTGATGATCACTGAAAGAC 720
Db 777 AATATTTTACTGTGTTTTCGAAGACTGAGTCTGTTGATCTCTGATGATCACTGAAAGAC 836
Qy 721 AATATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAA 780
Db 837 AATATCACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAA 896
Qy 781 GATAAATAGGCAATTCGGGCTCCAAACACTTGTGAAGTCCATTTTGAAACACCAAGATA 840
Db 897 GATAAATAGGCAATTCGGGCTCCAAACACTTGTGAAGTCCATTTTGAAACACCAAGATA 956
Qy 841 CTTGTGGAACACATCTCTGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC 900
Db 957 CTTGTGGAACACATCTCTGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC 1016
Qy 901 AACAGGGCCGGTTTCAGCATGGCAGCGTCTGGCTGGCTGCTCAAGAGATTGATTGAA 960
Db 1017 AACAGGGCCGGTTTCAGCATGGCAGCGTCTGGCTGGCTGCTCAAGAGATTGATTGAA 1076
Qy 961 ATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTGGA 1020
Db 1077 ATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAATTGGA 1136
Qy 1021 TTGATTCAGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTCATGGAGAGTATGACC 1080
Db 1137 TTGATTCAGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTCATGGAGAGTATGACC 1196
Qy 1081 TACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGCC 1140
Db 1197 TACCTCACAGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGCC 1256
Qy 1141 ATGTGTAGGTGTTACCTCGAGGCGCTGSCAGTGTGTGAGTGGGCTGCGAGATC 1200
Db 1257 ATGTGTAGGTGTTACCTCGAGGCGCTGSCAGTGTGTGAGTGGGCTGCGAGATC 1316
Qy 1201 CTCGGGGCTTGGCTACACAAGGAGTATCCGTACGAGCGCATCTGCTGTGACACCCG 1260
Db 1317 CTCGGGGCTTGGCTACACAAGGAGTATCCGTACGAGCGCATCTGCTGTGACACCCG 1376
Qy 1261 ATCTCTCATCTTCGAGGAGCAACATGAGATTCCTCGGATGTACATCGCCCTCACGGT 1320
Db 1377 ATCTCTCATCTTCGAGGAGCAACATGAGATTCCTCGGATGTACATCGCCCTCACGGT 1436
Qy 1321 CTCGAGCATGCGCGCCGATCCTGACTACCAAGGATCCATGAGCTTAAACAGGCAAAAGT 1380
Db 1437 CTCGAGCATGCGCGCCGATCCTGACTACCAAGGATCCATGAGCTTAAACAGGCAAAAGT 1496
Qy 1381 AGCAGATCATGATACCGTTGGCCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1440
Db 1497 AGCAGATCATGATACCGTTGGCCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1556
Qy 1441 CTGGGGCTGACAGCAACCACTGAGTTGTGACCCAGTCTTGGGAGAGTGGCCCAACAAG 1500
Db 1557 CTGGGGCTGACAGCAACCACTGAGTTGTGACCCAGTCTTGGGAGAGTGGCCCAACAAG 1616
Qy 1501 TTTGAGGAGAACACTTACTGCTTGGCCGGACCGTGGAGACACTGCTCCCGCTTTGGC 1560
Db 1617 TTTGAGGAGAACACTTACTGCTTGGCCGGACCGTGGAGACACTGCTCCCGCTTTGGC 1676
Qy 1561 AAGACCATCATGAGGAGCAGCTGTTGAGCGGTGGCCACATCCATCAACCTG 1620
Db 1677 AAGACCATCATGAGGAGCAGCTGTTGAGCGGTGGCCACATCCATCAACCTG 1736
Qy 1621 TATGCAATGAGCGGCTCTGCGGGCCAGCGCTCATCCGATTTGGGCTCCGCAAC 1680
Db 1737 TATGCAATGAGCGGCTCTGCGGGCCAGCGCTCATCCGATTTGGGCTCCGCAAC 1796
Qy 1681 CACGACCAAGAGGTTCTCTTGGCCCAACACCTTCTGCTGGAGGCTTACTTCGAATCTC 1740
Db 1797 CACGACCAAGAGGTTCTCTTGGCCCAACACCTTCTGCTGGAGGCTTACTTCGAATCTC 1856

Qy 1741 TTCAGCCTCTCTCAGTGGACAAGTATGCTCCAGAAAACCTAGATGACGACATTAAAGAA 1800
Db 1857 TTCAGCCTCTCTCAGTGGACAAGTATGCTCCAGAAAACCTAGATGACGACATTAAAGAA 1916
Qy 1801 GTGTCCAGCAGATCCTTTGAGAAGCGAGCCTATATCTGTGCCACCTCTGGACAGGACA 1860
Db 1917 GTGTCCAGCAGATCCTTTGAGAAGCGAGCCTATATCTGTGCCACCTCTGGACAGGACA 1976
Qy 1861 TGC 1863
Db 1977 TGC 1979
RESULT 3
ID ABV23498 standard; cDNA; 2522 BP.
XX ABV23498;
AC ABV23498;
XX 16-SEP-2002 (first entry)
DT Human prostate expression marker cDNA 23489.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
KW Homo sapiens.
OS XX
XX WO200160860-A2.
XX 23-AUG-2001.
PD 20-FEB-2001; 2001WO-US05171.
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
PI WPI; 2001-662795/76.
DR Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX Claim 1; Page 4299-4300; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;
SQ Query Match 100.0%; Score 1863; DB 23; Length 2522;

XX	PN	WO200160860-A2.	Db	477	TCCAAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCATCTGTG	536
XX	PD	23-AUG-2001.	Qy	421	ACCTTGGCAGCGCACCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG	480
XX	PF	20-FEB-2001; 2001WO-US05171.	Db	537	ACCTTGGCAGCGCACCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG	596
XX	PR	17-FEB-2000; 2000US-183319P.	Qy	481	CAGAAAGCCAAATACCTTGGCTTAAACTGGGCTCCGGGAGCAGATTTGAGGCTTCTGCCTC	540
XX	PR	16-MAR-2000; 2000US-189862P.	Db	597	CAGAAAGCCAAATACCTTGGCTTAAACTGGGCTCCGGGAGCAGATTTGAGGCTTCTGCCTC	656
XX	PR	25-MAY-2000; 2000US-207454P.	Qy	541	ACGAGAGCCAGCTGGGAGCGATGCGAGCTCAATCCGGAGCAGAGCCACACTAAAGTAA	600
XX	PR	09-JUN-2000; 2000US-211314P.	Db	657	ACGAGAGCCAGCTGGGAGCGATGCGAGCTCAATCCGGAGCAGAGCCACACTAAAGTAA	716
XX	PR	18-JUL-2000; 2000US-219007P.	Qy	601	GACAAAGACACTACATCTCAATGGCTCCAAAGTCTGGATTACTTAATGGAGGACTGGCC	660
XX	PR	13-DEC-2000; 2000US-255281P.	Db	717	GACAAAGACACTACATCTCAATGGCTCCAAAGTCTGGATTACTTAATGGAGGACTGGCC	776
XX	PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.	Qy	661	AAATTTTACTGTGTTTGCAGAAAGACTGAGTCTGGTTCGATTCGATGATGATGATGATG	720
XX	PI	Schlegel R, Endege WO, Monahan JE;	Db	777	AAATTTTACTGTGTTTGCAGAAAGACTGAGTCTGGTTCGATTCGATGATGATGATGATG	836
XX	XX	WPI; 2001-662795/76.	Qy	721	AAATTCACAGCAATTCATAGTAGAAAGAGACTTTGGTGAGTCACTTAATGGGAAACCCGAA	780
XX	XX	Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer	Db	837	AAATTCACAGCAATTCATAGTAGAAAGAGACTTTGGTGAGTCACTTAATGGGAAACCCGAA	896
XX	PS	Claim 1; Page 5464; 11750pp; English.	Qy	781	GATAAATAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGATA	840
XX	SS	The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:	Db	897	GATAAATAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGATA	956
XX	CC	(a) assessing whether a patient is afflicted with prostate cancer;	Qy	841	CTGTGGGAAACATCTTGGAGAGTTCGAGAGTGGTTCAGGTGCGCATGAACATCCTC	900
XX	CC	(b) monitoring the progression of prostate cancer in a patient;	Db	957	CTGTGGGAAACATCTTGGAGAGTTCGAGAGTGGTTCAGGTGCGCATGAACATCCTC	1016
XX	CC	(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;	Qy	901	AAACAGCGCGGCTTCAGCATGGGCGCTGCTGGTGGGCTGCTCAAGAGATTGATGAA	960
XX	CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;	Db	1017	AAACAGCGCGGCTTCAGCATGGGCGCTGCTGGTGGGCTGCTCAAGAGATTGATGAA	1076
XX	CC	(e) selecting a composition for inhibiting prostate cancer in a patient;	Qy	961	ATGACTGCTGAGTACGCCCTGCACAGGAAACAGTTTAAACAGAGGCTCAGTGAATTGGA	1020
XX	CC	(f) assessing the prostate cell carcinogenic potential of a compound;	Db	1077	ATGACTGCTGAGTACGCCCTGCACAGGAAACAGTTTAAACAGAGGCTCAGTGAATTGGA	1136
XX	CC	(g) determining whether prostate cancer has metastasized in a patient;	Qy	1021	TTGATTCAGGAGAAATTTGACCTGATGGCTCAGAGGCTTACGTCATGAGAGTATGACC	1080
XX	CC	(h) assessing the aggressiveness or indolence of prostate cancer in a patient;	Db	1137	TTGATTCAGGAGAAATTTGACCTGATGGCTCAGAGGCTTACGTCATGAGAGTATGACC	1196
XX	CC	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.	Qy	1081	TACCTCAGCAGCGGATGCTGGACCAACCTGGCTTCCGAGCTGCTCCATCGAGGAGCC	1140
XX	SQ	Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;	Db	1197	TACCTCAGCAGCGGATGCTGGACCAACCTGGCTTCCGAGCTGCTCCATCGAGGAGCC	1256
		Query Match 100.0%; Score 1863; DB 23; Length 2522;	Qy	1141	ATGGTGAAGGTGTTACGCTCCGAGGCGCTGCGAGTGTGTAGTGAAGGCTGACAGATC	1200
		Best Local Similarity 100.0%; Pred. No. 0;	Db	1257	ATGGTGAAGGTGTTACGCTCCGAGGCGCTGCGAGTGTGTAGTGAAGGCTGACAGATC	1316
		Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Qy	1201	CTCGGGGCTTGGGCTTACAAAGGACTATCCGTAGAGGCGATCTCGGTGACACCCGC	1260
Qy	1	ATGAGCGGCTGGGGCTCTTCTGCGCACACGAGCTGCGGCTCGCTGCTGCGGGGCTG	Db	1317	CTCGGGGCTTGGGCTTACAAAGGACTATCCGTAGAGGCGATCTCGGTGACACCCGC	1376
Db	117	ATGAGCGGCTGGGGCTCTTCTGCGCACACGAGCTGCGGCTCGCTGCTGCGGGGCTG	Qy	1261	ATCCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGT	1320
Qy	61	GTGGTCTTACCGCAACCGCGGCTACTGCGCACACGAGCCGCTGTACGAGCTTTCGCC	Db	1377	ATCCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGT	1436
Db	177	GTGGTCTTACCGCAACCGCGGCTACTGCGCACACGAGCCGCTGTACGAGCTTTCGCC	Qy	1321	CTGCAGCATCCGCGCCGATCTCTGACTACCAAGATCCATGAGTTCACAGGCTTAAACAGG	1380
Qy	121	AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGC	Db	1437	CTGCAGCATCCGCGCCGATCTCTGACTACCAAGATCCATGAGTTCACAGGCTTAAACAGG	1496
Db	237	AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGC	Qy	1381	AGCAGCTCATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGCAACTGTGGAC	1440
Qy	181	CAAGATGAATTAATGAATCAATCAGTTCTTGGGACCGCTGGGAAATTTCTTCACTGAA	Db	1497	AGCAGCTCATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGCAACTGTGGAC	1556
Db	297	CAAGATGAATTAATGAATCAATCAGTTCTTGGGACCGCTGGGAAATTTCTTCACTGAA	Qy	1441	CTGGGCTGACAGGCAACCATGGAGTTGTGACCCCATGTTTGGGAGCAGTGGCAACAG	1500
Qy	241	GAGGTGAGCTCCCGAAAAATGACCAAGAGGAAATCCAGATGAACACTTTGGAGAAA				
Db	357	GAGGTGAGCTCCCGAAAAATGACCAAGAGGAAATCCAGATGAACACTTTGGAGAAA				
Qy	301	TTGAAGAGCTTAGGCTTTTGGGCTGCAAGTCCCGAGAGATATGTTGGCTGGGCTTC				
Db	417	TTGAAGAGCTTAGGCTTTTGGGCTGCAAGTCCCGAGAGATATGTTGGCTGGGCTTC				
Qy	361	TCCAAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCATCTGTG				

Db 1557 CTGGGCTGACAGGCAACCATGGAGTTGTGACCCAGCTCTTGGGAGAGTCCCAACAG 1616
QY 1501 TTTGAGGAGAACACACTACTGCTTCGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1560
Db 1617 TTTGAGGAGAACACACTACTGCTTCGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1676
QY 1561 AAGACCATCATGGAGGAGCAGCTGCTGTAAGCGGTGGCCACATCTCTCATCAACCTG 1620
Db 1677 AAGACCATCATGGAGGAGCAGCTGCTGTAAGCGGTGGCCACATCTCTCATCAACCTG 1736
QY 1621 TATGCGCATGACGGCGCTGCTGCTGGGGCCAGCGCTCCATCCGCTATGGGCTCCGCAAC 1680
Db 1737 TATGCGCATGACGGCGCTGCTGCTGGGGCCAGCGCTCCATCCGCTATGGGCTCCGCAAC 1796
QY 1681 CACGACCAAGAGTCTCTTGGCCAAACACTTCTGCTGGAGCTTACTTGCAGAACTCTC 1740
Db 1797 CACGACCAAGAGTCTCTTGGCCAAACACTTCTGCTGGAGCTTACTTGCAGAACTCTC 1856
QY 1741 TTCAGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAGAAA 1800
Db 1857 TTCAGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAGAAA 1916
QY 1801 GTGTCCAGCAGATCTCTGAGAGCGAGCTATATCTGTGCCACCTCTGGACAGGACA 1860
Db 1917 GTGTCCAGCAGATCTCTGAGAGCGAGCTATATCTGTGCCACCTCTGGACAGGACA 1976
QY 1861 TGC 1863
Db 1977 TGC 1979

RESULT 5

ABV29355
ID ABV29355 standard; cDNA; 2522 BP.
AC ABV29355;
XX
DT 16-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 29346.
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
PN W0200160860-A2.
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 6287-6288; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 2522 BP; 642 A; 634 C; 701 G; 543 T; 2 other;

Query Match 100.0%; Score 1863; DB 23; Length 2522;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGGCTGCGGGCTCTTCTGCGCACACGCGCTGCGCTGCTGCGCTGCGGGGTCTG 60
Db 117 ATGAGCGGCTGCGGGCTCTTCTGCGCACACGCGCTGCGCTGCTGCGGGGTCTG 176
QY 61 GTGGTCTCTACCGCGAACCGGGGCTACTGCGCACACGCGCTGACGAGCTTTGCGC 120
Db 177 GTGGTCTCTACCGCGAACCGGGGCTACTGCGCACACGCGCTGACGAGCTTTGCGC 236
QY 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAAGTAGC 180
Db 237 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTCCTCCATTTCCAGAAGTAGC 296
QY 181 CAAGATGAATTAATGAATCAATCAGTTCTTGGGACCGCTGGAAATTTCTTCACTGAA 240
Db 297 CAAGATGAATTAATGAATCAATCAGTTCTTGGGACCGCTGGAAATTTCTTCACTGAA 356
QY 241 GAGGTGGACTCCCGAAATTTGACAGAGGAGGAAATCCAGATGAATTTGGAGAA 300
Db 357 GAGGTGGACTCCCGAAATTTGACAGAGGAGGAAATCCAGATGAATTTGGAGAA 416
QY 301 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGCGCTTGGGCTC 360
Db 417 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGCGCTTGGGCTC 476
QY 361 TCCAAACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGTGGTCCATCACTGTG 420
Db 477 TCCAAACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGTGGTCCATCACTGTG 536
QY 421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCCTGAGGAG 480
Db 537 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCCTGAGGAG 596
QY 481 CAGAAAGCCAAATACTTGCCTAACTGCGGGGAGCACATTCAGACCTTCTGCTTC 540
Db 597 CAGAAAGCCAAATACTTGCCTAACTGCGGGGAGCACATTCAGACCTTCTGCTTC 656
QY 541 ACGGAGCCAGCCAGTGGGCGGATGAGCTCAATCCGAGCAGAGCCACACTAGTGA 600
Db 657 ACGGAGCCAGCCAGTGGGCGGATGAGCTCAATCCGAGCAGAGCCACACTAGTGA 716
QY 601 GACAAGACACTTACATCTCAATGGCTCCAAAGTCTTGGATTTACTTAATGAGGACTGGCC 660
Db 717 GACAAGACACTTACATCTCAATGGCTCCAAAGTCTTGGATTTACTTAATGAGGACTGGCC 776
QY 661 AATATTTTACTGTGTTTGCAGAAAGTGAAGTCTGTTGATTTGATGATGATGATGATGAT 720
Db 777 AATATTTTACTGTGTTTGCAGAAAGTGAAGTCTGTTGATTTGATGATGATGATGATGAT 836
QY 721 AAATACACAGCATTCATAGTAGAAAGAGACTTTGTTGTTGAGTCACTAATGGGAAACCCGAA 780
Db 837 AAATACACAGCATTCATAGTAGAAAGAGACTTTGTTGTTGAGTCACTAATGGGAAACCCGAA 896
QY 781 GATAAATTAGGCATTCGGGGCTCCAAACACTTTGTGAAGTCCATTTTGTAAACACCAAGATA 840

Tue Sep 9 10:14:38 2003

us-09-945-326-3.rng

Db 897 GATAAATAGGCATTCGGGCTCCACACTTGTGAAGTCCATTTTGAACACCAAGATA 956
QY 841 CTTGTGGAACACATCTTGGAGGTCGAGAGTGGTTTAAGTGGCCATGAACATCTC 900
Db 957 CTTGTGGAACACATCTTGGAGGTCGAGAGTGGTTTAAGTGGCCATGAACATCTC 1016
QY 901 AACAGCGCGGTTTACAGTGGCCAGCGTCTGGTGGCTCTCAAGAGATTGATTGA 960
Db 1017 AACAGCGCGGTTTACAGTGGCCAGCGTCTGGTGGCTCTCAAGAGATTGATTGA 1076
QY 961 ATGACTGCTAGTACCGCTGCACAGGAACAGTTTAAAGAGGCTCAGTCAATTGGA 1020
Db 1077 ATGACTGCTAGTACCGCTGCACAGGAACAGTTTAAAGAGGCTCAGTCAATTGGA 1136
QY 1021 TTGATTTCAGGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTATGAGAGTATGACC 1080
Db 1137 TTGATTTCAGGAGAAATTTGACATGATGGCTCAGAAGGCTTACGTATGAGAGTATGACC 1196
QY 1081 TACCTCAGCAGGATGCTGGACCAACCTGCTTCCGACTGCTCCATCGAGGACGC 1140
Db 1197 TACCTCAGCAGGATGCTGGACCAACCTGCTTCCGACTGCTCCATCGAGGACGC 1256
QY 1141 ATGCTGAAGTGTTCAGCTCCGAGCGCCTGGCAGTGTGTGAGTGAGCGCTGCAGATC 1200
Db 1257 ATGCTGAAGTGTTCAGCTCCGAGCGCCTGGCAGTGTGTGAGTGAGCGCTGCAGATC 1316
QY 1201 CTCGGGGCTTGGCTACACAGGAGTATCCGTACAGCGCATCTCGTGACACCGC 1260
Db 1317 CTCGGGGCTTGGCTACACAGGAGTATCCGTACAGCGCATCTCGTGACACCGC 1376
QY 1261 ATCTCTCTCATCTTTCAGGGAACCAATGAGATTCCTCGGATGTACATCGCCCTGACGGT 1320
Db 1377 ATCTCTCTCATCTTTCAGGGAACCAATGAGATTCCTCGGATGTACATCGCCCTGACGGT 1436
QY 1321 CTCGAGCATCGGCGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAGTG 1380
Db 1437 CTCGAGCATCGGCGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAGTG 1496
QY 1381 AGCAGATCATGATACCTGTTGGCGGAGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1440
Db 1497 AGCAGATCATGATACCTGTTGGCGGAGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1556
QY 1441 CTGGGGCTCAGAGCAACATGGAGTTGTGACCCAGTCTTGGGACAGTGGCCAAACAG 1500
Db 1557 CTGGGGCTCAGAGCAACATGGAGTTGTGACCCAGTCTTGGGACAGTGGCCAAACAG 1616
QY 1501 TTTGAGGAGAACACCTACTGCTTTCGGCGGACCGTGGAGACACTGCTCGCTTGGC 1560
Db 1617 TTTGAGGAGAACACCTACTGCTTTCGGCGGACCGTGGAGACACTGCTCGCTTGGC 1676
QY 1561 AAGACCATCATGGAGGACAGTGTGTAAGCGGGTGGCAACATCTCATCAACCTG 1620
Db 1677 AAGACCATCATGGAGGACAGTGTGTAAGCGGGTGGCAACATCTCATCAACCTG 1736
QY 1621 TATGACATACGGCGTGTGTCGGGCGCAGCCGCTCATCCGATTTGGCTCGCAAC 1680
Db 1737 TATGACATACGGCGTGTGTCGGGCGCAGCCGCTCATCCGATTTGGCTCGCAAC 1796
QY 1681 CAGGACAGAGGTTCTTTCGCCAACACCTTCTCGGTGGAAAGCTTACTTGGAGATCTC 1740
Db 1797 CAGGACAGAGGTTCTTTCGCCAACACCTTCTCGGTGGAAAGCTTACTTGGAGATCTC 1856
QY 1741 TTCAGGCTCTCAGTGCACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAAGAA 1800
Db 1857 TTCAGGCTCTCAGTGCACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAAGAA 1916
QY 1801 GTGTCCAGCAGATCTTTCAGAGGAGGAGCTTATCTGTGCCACCTCTGACAGGACA 1860
Db 1917 GTGTCCAGCAGATCTTTCAGAGGAGGAGCTTATCTGTGCCACCTCTGACAGGACA 1976
QY 1861 TGC 1863
|||

Db 1977 TGC 1979
RESULT 6
AAS40878
ID AAS40878 standard; cDNA; 2492 BP.
XX
AC AAS40878;
DT 17-DEC-2001 (first entry)
XX cDNA encoding novel human enzyme polypeptide #94.
DE Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX Homo sapiens.
OS
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01239.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.

Db	247	CAAGATGAACCTTAATGAATCAATCAGTTCTTTGGGACCGGTGGAAAAATTCCTTCACTGAA	306
QY	241	GAGTGGACTCCCGAAAAATTGACAGGAAGGAAAAATCCAGATGAACACTTTGGAGAAA	300
Db	307	GAGTGGACTCCCGAAAAATTGACAGGAAGGAAAAATCCAGATGAACACTTTGGAGAAA	366
QY	301	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAAGAGATATGGTGGCCCTGGGCTTC	360
Db	367	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAAGAGATATGGTGGCCCTGGGCTTC	426
QY	361	TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCCATCACTGTG	420
Db	427	TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCCATCACTGTG	486
QY	421	ACCTGGCAGCGCACCAAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCTAGAGAG	480
Db	487	ACCTGGCAGCGCACCAAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCTAGAGAG	546
QY	481	CAGAAAGCCAAATACTTGGCTTAACCTGGGCTCCGGGAGCACATTCGAGCCTTCTGCCTC	540
Db	547	CAGAAAGCCAAATACTTGGCTTAACCTGGGCTCCGGGAGCACATTCGAGCCTTCTGCCTC	606
QY	541	ACGAGAGCAGCCAGTGGAGCGATGCAAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAA	600
Db	607	ACGAGAGCAGCCAGTGGAGCGATGCAAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAA	666
QY	601	GACAAGAGCACTACATCTTCAATGGCTCCCAAGGCTCTGATTAATGGAGGACTGGCC	660
Db	667	GACAAGAGCACTACATCTTCAATGGCTCCCAAGGCTCTGATTAATGGAGGACTGGCC	726
QY	661	AATATTTTACTGTGTTGCAAGACTGAGTCTGTTGATTTCTGATGGATCAGTGAAGAC	720
Db	727	AATATTTTACTGTGTTGCAAGACTGAGTCTGTTGATTTCTGATGGATCAGTGAAGAC	786
QY	721	AAATACAGCAATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGAGAACCCGAA	780
Db	787	AAATACAGCAATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGAGAACCCGAA	846
QY	781	GATAAATTAGGCATTGGGGCTCCAAACACTTTGTGAAGTCCATTTTGAACACCAAGATA	840
Db	847	GATAAATTAGGCATTGGGGCTCCAAACACTTTGTGAAGTCCATTTTGAACACCAAGATA	906
QY	841	CCTGTGGAACATCCTTTGGAGAGTGGAGATGGGTTTAGGTGGCCATCAACATCCTC	900
Db	907	CCTGTGGAACATCCTTTGGAGAGTGGAGATGGGTTTAGGTGGCCATCAACATCCTC	966
QY	901	ACAGCGCGGCTTCAGCATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAA	960
Db	967	ACAGCGCGGCTTCAGCATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAA	1026
QY	961	ATGACTGTGCTAGTACGCTGCAAGAGAAACAGTTTAAACAGAGCTCAGTGAATTGGA	1020
Db	1027	ATGACTGTGCTAGTACGCTGCAAGAGAAACAGTTTAAACAGAGCTCAGTGAATTGGA	1086
QY	1021	TTGATTCAGAGAAATTTGCACTGATGCTCAGAAGGCTTACGTATGAGAGATGATGACC	1080
Db	1087	TTGATTCAGAGAAATTTGCACTGATGCTCAGAAGGCTTACGTATGAGAGATGATGACC	1146
QY	1081	TACCTCACAGCAGGATGCTGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGCAGCC	1140
Db	1147	TACCTCACAGCAGGATGCTGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGCAGCC	1206
QY	1141	ATGGTGAAGGTGTTACGTCCAGGCGGCTGGCAGTGTGTCAGTGAAGCGCTGCAGATC	1200
Db	1207	ATGGTGAAGGTGTTACGTCCAGGCGGCTGGCAGTGTGTCAGTGAAGCGCTGCAGATC	1266
QY	1201	CTCGGGGCTTGGGCTACACAAGGGGACTATCGGTACGAGCGCATACTGCGTGACACCGC	1260
Db	1267	CTCGGGGCTTGGGCTACACAAGGGGACTATCGGTACGAGCGCATACTGCGTGACACCGC	1326
QY	1261	ATCCTCCTCATCTTCAGGGAACCAATGATGATTCCTCGGATGTACATCCCGCTGAGGGT	1320
Db	1327	ATCCTCCTCATCTTCAGGGAACCAATGATGATTCCTCGGATGTACATCCCGCTGAGGGT	1386
QY	1371	CTGCAGCATGCGCGCGCATCTCTGACTACAGAGTCCATGAGCTTAAACAGGCGCAAGTG	1380
Db	1387	CTGCAGCATGCGCGCGCATCTCTGACTACAGAGTCCATGAGCTTAAACAGGCGCAAGTG	1446
QY	1381	AGCACAGTCAATGATACCGTTGGCGCGGAGGCTTCGGGACTCCCTGGGCGGCAACTGTGAC	1440
Db	1447	AGCACAGTCAATGATACCGTTGGCGCGGAGGCTTCGGGACTCCCTGGGCGGCAACTGTGAC	1506
QY	1441	CTGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGTCTTGGGAGACTGTCGCTCCGCTTTGGC	1500
Db	1507	CTGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGTCTTGGGAGACTGTCGCTCCGCTTTGGC	1566
QY	1501	TTTGAGGAGAACACCTACTTGTGGCGGAGCCGCTGGGAGACTGTCGCTCCGCTTTGGC	1626
Db	1567	TTTGAGGAGAACACCTACTTGTGGCGGAGCCGCTGGGAGACTGTCGCTCCGCTTTGGC	1686
QY	1561	AAGACCATCATGGAGGAGCAGTGTGTAAGCGGGTGGCAACATCCTCATCAACCTG	1620
Db	1627	AAGACCATCATGGAGGAGCAGTGTGTAAGCGGGTGGCAACATCCTCATCAACCTG	1686
QY	1621	TATGGCATGACGCGCGTGTGTGCGGGCCAGCCGCTCCATCCGCAATTTGGGCTCCGCAAC	1680
Db	1687	TATGGCATGACGCGCGTGTGTGCGGGCCAGCCGCTCCATCCGCAATTTGGGCTCCGCAAC	1746
QY	1681	CACGACCGAGGTTCTCTTGGCCAAACACTTCTCGTGGAGGCTTACTTGCAGAACTCTC	1740
Db	1747	CACGACCGAGGTTCTCTTGGCCAAACACTTCTCGTGGAGGCTTACTTGCAGAACTCTC	1806
QY	1741	TTTCCGCTCTCTCAGCTGGAGCAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAGAAA	1800
Db	1807	TTTCCGCTCTCTCAGCTGGAGCAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAGAAA	1866
QY	1801	GTGTCGCCAGCATCTTTCAGAGGAGCCTATCTGTGCCACCTCTGTCAGAGGACA	1860
Db	1867	GTGTCGCCAGCATCTTTCAGAGGAGCCTATCTGTGCCACCTCTGTCAGAGGACA	1926
QY	1861	TGC 1863	
Db	1927	TGC 1929	
RESULT 7			
AAH24246			
ID	AAH24246 standard; cdna; 2440 BP.		
XX	AAH24246;		
AC	AAH24246;		
XX	11-SEP-2001 (first entry)		
DT	Human oxidoreductase protein ORP-24 encoding cdna.		
XX	Human oxidoreductase protein; ORP; cell proliferative disorder;		
KW	arteriosclerosis; cirrhosis; psoriasis; cancer; endocrine disorder;		
KW	diabetes mellitus; diabetes insipidus; dwarfism; hirsutism; amenorrhea;		
KW	osteoporosis; metabolic disorder; obesity; phenylketonuria;		
KW	hypercholesterolaemia; reproductive cycle defect; infertility;		
KW	ovulatory defect; menstrual cycle defect; endometriosis;		
KW	polycystic ovary disease; spermatogenesis disruption; impotence;		
KW	neurological disorder; epilepsy; stroke; Alzheimer's disease;		
KW	Huntington's disease; Parkinson's disease; Creutzfeldt-Jakob disease;		
KW	meningitis; cerebral palsy; muscular dystrophy; mood disorder; anxiety;		
KW	schizophrenic disorder; infection; autoimmune disorder;		
KW	inflammatory disorder; acquired immunodeficiency syndrome; AIDS; asthma;		
KW	allergy; Crohn's disease; atopic dermatitis; gout; multiple sclerosis;		
KW	rheumatoid arthritis; ulcerative colitis; drug screening;		
KW	toxicity screening; transgenic animal; SNP detection; gene therapy; ss.		
OS	Homo sapiens.		
XX	Key		
PH	Location/Qualifiers		
FT	60..1925		
FT	/*tag= a		

FT

XX /product= "ORP-24"

XX /function= "Oxidoreductase"

PN

XX WO200144448-A2.

XX

PD 21-JUN-2001.

XX

XX 07-DEC-2000; 2000WO-US33158.

XX

XX 16-DEC-1999; 99US-0172367.

XX

XX (INCY-) INCYTE GENOMICS INC.

XX

XX Yue H, Lal P, Tang YT, Hillman JL, Baughn MR, Azimzai Y, Lu DAM;

XX

XX WPI: 2001-390245/41.

XX

XX P-PSDB; AAB73691.

XX

XX Novel human oxidoreductase protein (ORP) useful for diagnosing, treating and preventing cell proliferative, neurological, viral, reproductive and autoimmune/inflammatory disorders associated with abnormal expression of ORP

XX

XX Claim 5; Page 134; 136pp; English.

XX

Sequences AAB73668-AAB73694 represent 27 novel human oxidoreductase proteins, designated ORP-1 to ORP-27 respectively, and sequences AAH24223-AAH24249 represent cDNAs encoding ORP-1 to ORP-27. Human ORP proteins and nucleic acids are useful for diagnosing, treating or preventing cell proliferative disorders (e.g. arteriosclerosis, cirrhosis, psoriasis, cancers); endocrine disorders (e.g. diabetes mellitus, diabetes insipidus, dwarfism, hirsutism, amenorrhea, osteoporosis); metabolic disorders (e.g., obesity, phenylketonuria, hypercholesterolemia); reproductive disorders (e.g., infertility, ovulatory and menstrual cycle defects, endometriosis, polycystic ovary disease, disruption of spermatogenesis, impotence); neurological disorders (e.g., epilepsy, stroke, Alzheimer's disease, Huntington's disease, Parkinson's disease, meningitis, Creutzfeldt-Jakob disease, cerebral palsy, muscular dystrophy, mood, anxiety and schizophrenic disorders); viral, bacterial, fungal and parasitic infections; and autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome (AIDS), allergies, asthma, Crohn's disease, atopic dermatitis, gout, multiple sclerosis, rheumatoid arthritis or ulcerative colitis. Human ORP proteins and nucleotides can be used to identify compounds which modulate their activity or expression. ORP nucleic acid sequences may also be used for assessing the toxicity of a test compound, to detect upstream sequences such as promoters and regulatory elements, and to create knock out or knock in animals or transgenic animals to model human disease. Oligonucleotide primers derived from ORP gene sequences may be used to detect single nucleotide polymorphisms (SNPs) and for mapping the naturally occurring genomic sequences. Antibodies specific for ORP proteins may be used in the diagnosis of disorders associated with aberrant ORP expression, in assays to monitor patients being treated with ORP or modulators thereof, and for assessing toxicity of potential drugs.

XX Sequence 2440 BP; 611 A; 614 C; 682 G; 533 T; 0 other;

XX

XX Query Match

XX Best Local Similarity 99.98; Score 1861.4; DB 22; Length 2440;

XX

XX Matches 1862; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX

XX 1 ATGAGCGGCTCGCGGCTCTTCCTCGGCACACCGCTCGGCTCGTGCCTGCGCGGCTCG 60

XX

XX 60 ATGAGCGGCTCGCGGCTCTTCCTCGGCACACCGCTCGGCTCGTGCCTGCGCGGCTCG 119

XX

XX 61 GTGGTCTCTACCGGAAACCGCGGCTACTGCGCACCGCGGCTGTACGAGCTTCGCG 120

XX

XX 120 GTGGTCTCTACCGGAAACCGCGGCTACTGCGCACCGCGGCTGTACGAGCTTCGCG 179

XX

XX 121 AAAGAGCTTTTCTTAGCGAAATCAAGAAGAAGAGTTTCCCATTTCCAGAGTTAGC 180

XX

XX 180 AAAGAGCTTTTCTTAGCGAAATCAAGAAGAAGAGTTTCCCATTTCCAGAGTTAGC 239

XX

181 CAAGATGAATTAATGAATCAATCAGTTCTTGGACCCGTTGGAAAAATTTCTTCACTGAA 240
 240 CAAGATGAATTAATGAATCAATCAGTTCTTGGACCCGTTGGAAAAATTTCTTCACTGAA 299
 241 GAGGTGGACTCCCGAAAAATTCACAGGAAGGAAAAATCCAGATGAACATTTGGAGAAA 300
 300 GAGGTGGACTCCCGAAAAATTCACAGGAAGGAAAAATCCAGATGAACATTTGGAGAAA 359
 301 TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCACAGAAAGATATGGTGGCTGGGCTTC 360
 360 TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCACAGAAAGATATGGTGGCTGGGCTTC 419
 361 TCCAAACACCATGTACTCAAGACTAGGGAGAGATCATCAGCATGGATGGGTCCATCCTG 420
 420 TCCAAACACCATGTACTCAAGACTAGGGAGAGATCATCAGCATGGATGGGTCCATCCTG 479
 421 ACCCTGGCAGCGCACCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGCTGAGAG 480
 480 ACCCTGGCAGCGCACCAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGCTGAGAG 539
 481 CAGAAAGCCAAATATCTTGCCTAAACTGGCGTCCCGGGGAGCACATTGCAGGCTTCTGCTC 540
 540 CAGAAAGCCAAATATCTTGCCTAAACTGGCGTCCCGGGGAGCACATTGCAGGCTTCTGCTC 599
 541 ACGGAGCCAGCCAGTGGGCGGATGCGACCTCAATCCGGAGCAGAGCCACACTAAAGTAA 600
 600 ACGGAGCCAGCCAGTGGGCGGATGCGACCTCAATCCGGAGCAGAGCCACACTAAAGTAA 659
 601 GACAAGAGCACTACATCTTCAATGGCTCCAAAGTCTGGATTAATGAGGAGCTGGCC 660
 660 GACAAGAGCACTACATCTTCAATGGCTCCAAAGTCTGGATTAATGAGGAGCTGGCC 719
 661 AATATTTTACTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 720
 720 AATATTTTACTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 779
 721 AATATCACACATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGAAACCCGAA 780
 780 AATATCACACATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGAAACCCGAA 839
 781 GATAAATAGGATTCGGGCTCCAACTTTGTAAGTCCATTTTGAACACCAAGATA 840
 840 GATAAATAGGATTCGGGCTCCAACTTTGTAAGTCCATTTTGAACACCAAGATA 899
 841 CCTGTGAAACATCTTGGAGAGTTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC 900
 900 CCTGTGAAACATCTTGGAGAGTTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC 959
 901 AACAGCGCGGTTTCAAGCATGGGACGCTCGTGGCTGGCTCAAGAGATTGATTGAA 960
 960 AACAGCGCGGTTTCAAGCATGGGACGCTCGTGGCTGGCTCAAGAGATTGATTGAA 1019
 961 ATGACTGCTGAGTACGCTTCACAGGAAACAGTTTAAACAGAGGCTCAGTGAATTTGA 1020
 1020 ATGACTGCTGAGTACGCTTCACAGGAAACAGTTTAAACAGAGGCTCAGTGAATTTGA 1079
 1021 TTGATTACAGGAAATTTGCTGCTCAGAGGCTTACGTATGAGAGGATGATGACC 1080
 1080 TTGATTACAGGAAATTTGCTGCTCAGAGGCTTACGTATGAGAGGATGATGACC 1139
 1081 TACCTCACAGCAGGATGCTGACCAACCTGGCTTTCCGACTGCTCCATCGAGCAGCC 1140
 1140 TACCTCACAGCAGGATGCTGACCAACCTGGCTTTCCGACTGCTCCATCGAGCAGCC 1199
 1141 ATGGTGAAGGTTTTCAGCTCCGAGCGCTGGCAGTGTGTGAGTGGCGCTGCAGATC 1200
 1200 ATGGTGAAGGTTTTCAGCTCCGAGCGCTGGCAGTGTGTGAGTGGCGCTGCAGATC 1259
 1201 CTCGGGGCTTTGGGCTACACAGAGGACTATCCGTACGAGCGCATATCGTGGTACACCCG 1260
 1260 CTCGGGGCTTTGGGCTACACAGAGGACTATCCGTACGAGCGCATATCGTGGTACACCCG 1319

1261 ATCTCTCTCATCTTCGAGGGAACCAATGAGATTTCTCCGGATGTACATCGCCCTGACGGGT 1320
|||||
1320 ATCTCTCTCATCTTCGAGGGAACCAATGAGATTTCTCCGGATGTACATCGCCCTGACGGGT 1379
|||||
1321 CTGCAGCATGCGCGCGCATCTGACTACCGAGATCCATGAGCTTAAACAGGCCAAAGTG 1380
|||||
1380 CTGCAGCATGCGCGCGCATCTGACTACCGAGATCCATGAGCTTAAACAGGCCAAAGTG 1439
|||||
1381 AGCAGCATGATGATATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAACTGTGGAC 1440
|||||
1440 AGCAGCATGATGATATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGGAACTGTGGAC 1499
|||||
1441 CTGGGGCTACAGGCAACCATGAGTTGTGCACCCAGTCTTGGCGACAGTGCCCAACAG 1500
|||||
1500 CTGGGGCTACAGGCAACCATGAGTTGTGCACCCAGTCTTGGCGACAGTGCCCAACAG 1559
|||||
1501 TTTGAGGAGAACACCTACTGCTTCGGCGGAGCCGCTGGAGACACTGCTGCTCCGCTTTGGC 1560
|||||
1560 TTTGAGGAGAACACCTACTGCTTCGGCGGAGCCGCTGGAGACACTGCTGCTCCGCTTTGGC 1619
|||||
1561 AGACCATATGAGGAGCAGCTGTTGTAAGCGGCTGAGCGGAGTGGCAACATCTCATCAACCTG 1620
|||||
1620 AGACCATATGAGGAGCAGCTGTTGTAAGCGGCTGAGCGGAGTGGCAACATCTCATCAACCTG 1679
|||||
1621 TATGCGCATGACGCGCGCTGCTGTCGCGGGCCAGCCGCTCCATCCGATTTGGGCTCCGCAAC 1680
|||||
1680 TATGCGCATGACGCGCGCTGCTGTCGCGGGCCAGCCGCTCCATCCGATTTGGGCTCCGCAAC 1739
|||||
1681 CAGCAGCAGAGGTTCTTGTGGCCACACCTTCTGCGTGAGAGCTTCTGAGAGTCTGAGAGTCTC 1740
|||||
1740 CAGCAGCAGAGGTTCTTGTGGCCACACCTTCTGCGTGAGAGCTTCTGAGAGTCTGAGAGTCTC 1799
|||||
1741 TTAGCCTCTCTCAGCTGGACAAAGTATGCTTCCAGAAACCTAGATGAGCAGATTAAGAAA 1800
|||||
1800 TTAGCCTCTCTCAGCTGGACAAAGTATGCTTCCAGAAACCTAGATGAGCAGATTAAGAAA 1859
|||||
1801 GTGTCCAGCAGATCTCTGAGAGCGAGCGCTATATCTGTGCCACCCCTCTGGCAGGACA 1860
|||||
1860 GTGTCCAGCAGATCTCTGAGAGCGAGCGCTATATCTGTGCCACCCCTCTGGCAGGACA 1919
|||||
1861 TGC 1863
|||
1920 TGC 1922
|||

RESULT 8
AAH16781
ID AAH16781 standard; cDNA; 2399 BP.
XX
AC AAH16781;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:16009.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI; 2001-318749/34.
Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
and/or diagnosis of the abnormality of the proteins encoded by the
full-length cDNAs -
Claim 8; SEQ ID 16009; 2537pp + CD ROM; English.
The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dr primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
oligonucleotide comprises at least 15 nucleotides and the combination of
the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.
Sequence 2399 BP; 597 A; 606 C; 672 G; 524 T; 0 other;
Query Match 99.8%; Score 1859.8; DB 22; Length 2399;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1861; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGCGGCTCGGGGCTCTTCTGCGCACCAACCGCGGCTGCTGCGCACCAACCGCGGCTG 60
DB 47 ATGAGCGGCTCGGGGCTCTTCTGCGCACCAACCGCGGCTGCTGCGCACCAACCGCGGCTG 106
QY 61 GTGGTCTCTACCGCGAAACCGCGGCTGCTGCGCACCAACCGCGGCTGCTGCGCACCAAC 120
DB 107 GTGGTCTCTACCGCGAAACCGCGGCTGCTGCGCACCAACCGCGGCTGCTGCGCACCAAC 166
QY 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGAGTTTCCCATTTCCAGAGTTAGC 180
DB 167 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGAGTTTCCCATTTCCAGAGTTAGC 226
QY 181 CAAGATGAACCTTAATGAATCAATCAATGTTCTGGACCCGCTGAGAAATTTCTCACTGAA 240
DB 227 CAAGATGAACCTTAATGAATCAATCAATGTTCTGGACCCGCTGAGAAATTTCTCACTGAA 286
QY 241 GAGGTGAGCTCCCGAAAAATTTGACAGGAGGAGAAATCCAGATGAAACTTTTGAGAAA 300
DB 287 GAGGTGAGCTCCCGAAAAATTTGACAGGAGGAGAAATCCAGATGAAACTTTTGAGAAA 346
QY 301 TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGGGCTTC 360
DB 347 TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGGGCTTC 406
QY 361 TCCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTG 420
DB 407 TCCAACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGTCCATCACTGTG 466
QY 421 ACCCTGGCAGCGCACCAAGGCTATTGGCTCAAGGGATCATCTTGGCTGCACTGAGGAG 480
DB 467 ACCCTGGCAGCGCACCAAGGCTATTGGCTCAAGGGATCATCTTGGCTGCACTGAGGAG 526

PS	Claim 5; Page 2343-2344; 5507pb; English.	
XX	AACT74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulneryary; antiparasitic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticonvulsant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic erythematous, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive.	
XX	SQ	Sequence 2218 BP; 573 A; 548 C; 609 G; 487 T; 1 other;
	Query Match	90.8%; Score 1690.8; DB 21; Length 2218;
	Best Local Similarity	99.9%; Pred. No. 0;
	Matches 1692; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	170	CAGAAGTTAGCCAAAGATGAATTAATGAATCAATCAGTTCTTGGACCCGCGTGGAAAAAT 229
DB	2	CAGAGCTAGCCAAAGATGAATTAATGAATCAATCAGTTCTTGGACCCGCGTGGAAAAAT 61
QY	230	TCCTACTGAAGAGTGGACTCCGAAAAATTGACAGAGAGGAAAAATCCAGATGAAA 289
DB	62	TCCTACTGAAGAGTGGACTCCGAAAAATTGACAGAGAGGAAAAATCCAGATGAAA 121
QY	290	CTTTGGAGAAATTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTG 349
DB	122	CTTTGGAGAAATTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTG 181
QY	350	GCCTGGGCTTCTCCACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGGT 409
DB	182	GCCTGGGCTTCTCCACACCATGTACTCAAGACTAGGGAGATCATCAGCATGGATGGGT 241
QY	410	CCATCACTGTGACCTGCGAGCCGACCAAGCTATTGGCTCAAGGGATCATCTTGGCTG 469
DB	242	CCATCACTGTGACCTGCGAGCCGACCAAGCTATTGGCTCAAGGGATCATCTTGGCTG 301
QY	470	GCATGAGGAGCAGAAAGCCAAATACCTTAACTGGCTCGGGGAGCAGATTCGAG 529
DB	302	GCATGAGGAGCAGAAAGCCAAATACCTTAACTGGCTCGGGGAGCAGATTCGAG 361
QY	530	CTTCTGCTCAGCGAGCCAGCTAGTGGGAGGATGAGCTCAATCCGAGCAGAGCCA 589
DB	362	CTTCTGCTCAGCGAGCCAGCTAGTGGGAGGATGAGCTCAATCCGAGCAGAGCCA 421
QY	590	CACCTAAGTGAAGCAAGAAGCAGCTACATCTCAATGGCTCCAAAGTCTGGATTACTAATG 649
DB	422	CACCTAAGTGAAGCAAGAAGCAGCTACATCTCAATGGCTCCAAAGTCTGGATTACTAATG 481
QY	650	GAGGACTGGCCAAATATTTTACTGTGTTGCAAGAGCTAGGTCGTTGATCTCATGAT 709
DB	482	GAGGACTGGCCAAATATTTTACTGTGTTGCAAGAGCTAGGTCGTTGATCTCATGAT 541
QY	710	CAGTGAAGACAAATCAGCATTATAGTAGAAGAGACTTTTGGTGGAGTCACTAATG 769
DB	542	CAGTGAAGACAAATCAGCATTATAGTAGAAGAGACTTTTGGTGGAGTCACTAATG 601
QY	770	GGAAACCCGAAAGATAAATTAGGCATTCGGGCTCCAAACACTTGTGAAGTCCATTTGAAA 829
DB	602	GGAAACCCGAAAGATAAATTAGGCATTCGGGCTCCAAACACTTGTGAAGTCCATTTGAAA 661
QY	830	ACACCAAGATACCTGTGGAAAAACATCCTTGGAGAGTTCGGAGATGGGTTTAAAGTGGCCA 889
DB	662	ACACCAAGATACCTGTGGAAAAACATCCTTGGAGAGTTCGGAGATGGGTTTAAAGTGGCCA 721
QY	890	TGAACATCCTCAACAGCGCGGTTTCACATCGGAGCGTGTGGCTGGGCTGGCTCAAGA 949
DB	722	TGAACATCCTCAACAGCGCGGTTTCACATCGGAGCGTGTGGCTGGGCTGGCTCAAGA 781
QY	950	GATTGATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGGCTCA 1009
DB	782	GATTGATTGAATGACTGCTGAGTACGCTGCACAAAGAAACAGTTTAAACAGAGGCTCA 841
QY	1010	GTGAATTTGGATTGATTGAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGG 1069
DB	842	GTGAATTTGGATTGATTGAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTCATGG 901
QY	1070	AGAGTATGACCTACCTCACAGCAGGATGCTGGACCAACCTGGCTTCCCGACTGCTCCA 1129
DB	902	AGAGTATGACCTACCTCACAGCAGGATGCTGGACCAACCTGGCTTCCCGACTGCTCCA 961
QY	1130	TCGAGGAGCCCATGTTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGG 1189
DB	962	TCGAGGAGCCCATGTTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGG 1021
QY	1190	CGCTGAGATCCTCGGGGCTTGGGCTACACAAAGGAGCTATCCGTACGAGCGCATCTGC 1249
DB	1022	CGCTGAGATCCTCGGGGCTTGGGCTACACAAAGGAGCTATCCGTACGAGCGCATCTGC 1081
QY	1250	GTGACACCCGATCCTCCCTCATCTTTCGAGGAGAAACAAATGAGATTTCTCCGGATGTACATCG 1309
DB	1082	GTGACACCCGATCCTCCCTCATCTTTCGAGGAGAAACAAATGAGATTTCTCCGGATGTACATCG 1141
QY	1310	CCCTGACGGTCTGACAGCATGCCGCCGCATCCTGACTACACAGGATCCATGAGCTTAAAC 1369
DB	1142	CCCTGACGGTCTGACAGCATGCCGCCGCATCCTGACTACACAGGATCCATGAGCTTAAAC 1201
QY	1370	AGGCCAAAGTGAACAGCATGATGATACCGTTCGCGGAGGCTTCGGGACTCCCTGGGCG 1429
DB	1202	AGGCCAAAGTGAACAGCATGATGATACCGTTCGCGGAGGCTTCGGGACTCCCTGGGCG 1261
QY	1430	GAACTGTGACCTGGGGCTGACAGGCAACCATGAGTTGTGACACCCAGTCTTTCGCGGACA 1489
DB	1262	GAACTGTGACCTGGGGCTGACAGGCAACCATGAGTTGTGACACCCAGTCTTTCGCGGACA 1321
QY	1490	GTGCCAAAGTTTGAAGAGACACACTACTGCTTCGCGCGGACCTGGAGACACTGCTGC 1549
DB	1322	GTGCCAAAGTTTGAAGAGACACACTACTGCTTCGCGCGGACCTGGAGACACTGCTGC 1381
QY	1550	TCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGTTGCTGAAAGCGGGTGGCCAAATCC 1609
DB	1382	TCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGTTGCTGAAAGCGGGTGGCCAAATCC 1441
QY	1610	TCATCAACCTGTATGCGATGACGGCGTGTGTCGGGGGCGAGCGCTCCATCCGATTTG 1669
DB	1442	TCATCAACCTGTATGCGATGACGGCGTGTGTCGGGGGCGAGCGCTCCATCCGATTTG 1501
QY	1670	GGCTCCGCAACAGCCAGCCAGAGTCTCTTTCGCGCAACACTTCTTCGCTGGAAGCTTACT 1729
DB	1502	GGCTCCGCAACAGCCAGCCAGAGTCTCTTTCGCGCAACACTTCTTCGCTGGAAGCTTACT 1561
QY	1730	TGCAGATCTCTTCAGCTCTCTCAGCTGGACAAAGTATGTCGAGAAACCTAGATGAGC 1789
DB	1562	TGCAGATCTCTTCAGCTCTCTCAGCTGGACAAAGTATGTCGAGAAACCTAGATGAGC 1621
QY	1790	AGATTAGAAGTGTCCACAGCATCTTTCAGAGAGCAGCTATATCTGTGCCACCCCTC 1849
DB	1622	AGATTAGAAGTGTCCACAGCATCTTTCAGAGAGCAGCTATATCTGTGCCACCCCTC 1681
QY	1850	TGGACAGGACATGC 1863
DB	1682	TGGACAGGACATGC 1695

RESULT 10
AAH15804
ID AAH15804 standard; cDNA; 2400 BP.
AC AAH15804;
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:14270.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EPI074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Salto K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 14270; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95983 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 2400 BP; 570 A; 629 C; 685 G; 516 T; 0 other;
Query Match 89.7%; Score 1671.4; DB 22; Length 2400;
Best Local Similarity 92.9%; Pred. No. 0;
Matches 1807; Conservative 0; Mismatches 56; Indels 82; Gaps 2;
Oy 1 ATGAGCGGCTCGGGCTCTTCCTGCGCACACCGCTCGGCTCGTGCCTCCGGGCTCTG 60
Db 47 ATGAGCGGCTCGGGCTCTTCCTGCGCACACCGCTCGGCTCGTGCCTCCGGGCTCTG 106

Qy	61	GTGGTCTCTACCGGAACCGCGGCTACTGCGCACAGCCCGCTGTACGAGCTTTCGCC	120
Db	107	GTGGTCTCTACCGGAACCGCGGCTACTGCGCACAGCCCGCTGTACGAGCTTTCGCC	166
Qy	121	AAGAGCTTTTCTAGGCAAAATCAAGAAAGAA-----	153
Db	167	AAGAGCTTTTCTAGGCAAAATCAAGAAAGTAAACGAGAGCCCTGGGGAACCCCTTGCTG	226
Qy	154	-----CAAGTTCCTCCATTTCCAG	172
Db	227	TCGTGGTCCCGCTTTTCCACCTCAGCTGCAAGAGCTGGTGTGAAGCTTTGTGAGATTCGCC	286
Qy	173	AAGTTAGCCCAAGATGA-----	218
Db	287	AAACCTGCCAGAGAGATACACCCCTCGCGCCGAGGCGTGTAAACACTCTCGGATTCCTGAGT	346
Qy	219	CGTGGAAAAATTCCTACTGAAGAGCTGACTCCCGAAAAATTTACCAAGAGGAAAAAT	278
Db	347	TCCAGGAAACCTTCCAGAGAGAAAGTGGACTCCCGAAAAATTTACCAAGAGGAAAAAT	406
Qy	279	CCAGATGAAACTTTTGAGAGAAATTCAGAGAGCTAGGCGCTTTTGGGCTGCAAGTCCCGAGA	338
Db	407	CCAGATGAAACTTTTGAGAGAAATTCAGAGAGCTAGGCGCTTTTGGGCTGCAAGTCCCGAGA	466
Qy	339	AGATATGCTGGCGCTGGCTTCTCCACACCATGTACTCAAGACTAGGGGAGATCATCAG	398
Db	467	AGAATATGCTGGCGCTGGCTTCTCCACACCATGTACTCAAGACTAGGGGAGATCATCAG	526
Qy	399	CATGATGGTCCATCACTGTGACCCCTGGCAGCCAGCCAGGCTATTGGCCCTCAAGGGAT	458
Db	527	CATGATGGTCCATCACTGTGACCCCTGGCAGCCAGCCAGGCTATTGGCCCTCAAGGGAT	586
Qy	459	CATCTGGCTGGCCTGAGGAGAGAAAGCCAAATACTTGGCTAAACTGGCGTCCGGGA	518
Db	587	CATCTGGCTGGCCTGAGGAGAGAAAGCCAAATACTTGGCTAAACTGGCGTCCGGGA	646
Qy	519	GCACATGACAGCCCTTCTGCTCAGCGGAGCCAGCCAGTGGGAGCGATCAAGTCCG	578
Db	647	GCACATGACAGCCCTTCTGCTCAGCGGAGCCAGCCAGTGGGAGCGATCAAGTCCG	706
Qy	579	GAGCAGAGCCACACTAAGTGAAGACAAAGACACACTACATCTCAATGGCTCCAAGTCTG	638
Db	707	GAGCAGAGCCACACTAAGTGAAGACAAAGACACACTACATCTCAATGGCTCCAAGTCTG	766
Qy	639	GATTACTAATGGAGGACTGGCCATATTTTACTGTGTTTGCAGAAAGACTGAGGTCGTTGA	698
Db	767	GATTACTAATGGAGGACTGGCCATATTTTACTGTGTTTGCAGAAAGACTGAGGTCGTTGA	826
Qy	699	TTCTGTATGGATCAGTGAAGACAAATCAGAGCATTCATAGTAGAAGAGACATTTGGTGG	758
Db	827	TTCTGTATGGATCAGTGAAGACAAATCAGAGCATTCATAGTAGAAGAGACATTTGGTGG	886
Qy	759	AGTCACTAATGGGAAACCCGAAAGATAAATTTAGGCATTTCCGGGCTCCCAACACTTGTGAAGT	818
Db	887	AGTCACTAATGGGAAACCCGAAAGATAAATTTAGGCATTTCCGGGCTCCCAACACTTGTGAAGT	946
Qy	819	CCATTTTGAACACCAAGATACCTGTGGAAACATCTTGGAGAGGTCGAGATGGGTT	878
Db	947	CCATTTTGAACACCAAGATACCTGTGGAAACATCTTGGAGAGGTCGAGATGGGTT	1006
Qy	879	TAAGTGGCCATGAACATCCTCAACAGCGCCGGTTTTCAGCATGGGAGGCTCGTGGCTGG	938
Db	1007	TAAGTGGCCATGAACATCCTCAACAGCGCCGGTTTTCAGCATGGGAGGCTCGTGGCTGG	1066
Qy	939	GCTGCTCAGAGATTTGATTAATGACTGCTGAGTACGCTGCACAGGAAACAGTTTAA	998
Db	1067	GCTGCTCAGAGATTTGATTAATGACTGCTGAGTACGCTGCACAGGAAACAGTTTAA	1126
Qy	999	CAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACTCATGGCTCAGAGGC	1058
Db	1127	CAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACTCATGGCTCAGAGGC	1186

Tue Sep 9 10:14:38 2003

us-09-945-326-3.rng

QY 1059. TTACGTCATGAGAGTATGACCTACACAGCAGGATGCTGGACCAACCTGGCTTCC 1118
DB |||||||
QY 1187 TTACGTCATGAGAGTATGACCTACACAGCAGGATGCTGGACCAACCTGGCTTCC 1246
DB |||||||
QY 1119 CGACTGCTCCATCGAGGCGCCATGCTGAAGGTTCAGTCCGAGGCGCTGGCAGTG 1178
DB |||||||
QY 1247 CGACTGCTCCATCGAGGCGCCATGCTGAAGGTTCAGTCCGAGGCGCTGGCAGTG 1306
DB |||||||
QY 1179 TGTGAGTAGGCGCTGCAGATCCTCGGGGGCTGGCTACACAAGGACTATCCGTAGA 1238
DB |||||||
QY 1307 TGTGAGTAGGCGCTGCAGATCCTCGGGGGCTGGCTACACAAGGACTATCCGTAGA 1366
DB |||||||
QY 1239 GGCATACATGCGTGCAGCAGGCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCG 1298
DB |||||||
QY 1367 GGCATACATGCGTGCAGCAGGCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCG 1426
DB |||||||
QY 1299 GATGTATATGCGCCTGACGGGTCTGAGATGCGCGCGCATCTGACTACAGATCCA 1358
DB |||||||
QY 1427 GATGTATATGCGCCTGACGGGTCTGAGATGCGCGCGCATCTGACTACAGATCCA 1486
DB |||||||
QY 1359 TGAGCTTAACAGGCCAAAGTGAGCACAGTTCATGATACGTTGGCGGAGGCTTCGGGA 1418
DB |||||||
QY 1487 TGAGCTTAACAGGCCAAAGTGAGCACAGTTCATGATACGTTGGCGGAGGCTTCGGGA 1546
DB |||||||
QY 1419 CTCCTGGGCGCAACTGTGACCTGGGCTGACAGGCAACCATGGAGTTGTGACCCAG 1478
DB |||||||
QY 1547 CTCCTGGGCGCAACTGTGACCTGGGCTGACAGGCAACCATGGAGTTGTGACCCAG 1606
DB |||||||
QY 1479 TCTTGGGACAGTGCACCAAGTTTGAGGAGACACCTACTGCTTCGGCGGACCGTGA 1538
DB |||||||
QY 1607 TCTTGGGACAGTGCACCAAGTTTGAGGAGACACCTACTGCTTCGGCGGACCGTGA 1666
DB |||||||
QY 1539 GACACTGCTCTCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGT 1598
DB |||||||
QY 1667 GACACTGCTCTCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGT 1726
DB |||||||
QY 1599 GGCCAAACATCCTATCAACCTGTATGGCATGACGGCGGTCTGTCGGGCGGACCGCTC 1658
DB |||||||
QY 1727 GGCCAAACATCCTATCAACCTGTATGGCATGACGGCGGTCTGTCGGGCGGACCGCTC 1786
DB |||||||
QY 1659 CATCCGATTTGGGCTCGCAACACACAGGCTTCTTGGCCAAACACCTTCGCGT 1718
DB |||||||
QY 1787 CATCCGATTTGGGCTCGCAACACACAGGCTTCTTGGCCAAACACCTTCGCGT 1846
DB |||||||
QY 1719 GGAAGCTTACTTGAGAACTCTTTCAGCTCTCTGAGTGGCAAGTATGCTCCAGAAA 1778
DB |||||||
QY 1847 GGAAGCTTACTTGAGAACTCTTTCAGCTCTCTGAGTGGCAAGTATGCTCCAGAAA 1906
DB |||||||
QY 1779 CCTAGATGACAGATTAAAGAGTGTCCAGCAGATCCTTTGAGAAGCGAGCTATATCTG 1838
DB |||||||
QY 1907 CTTAGATGACAGATTAAAGAGTGTCCAGCAGATCCTTTGAGAAGCGAGCTATATCTG 1966
DB |||||||
QY 1839 TGCCACCCCTCTGGACAGGACATGC 1863
DB |||||||
QY 1967 TGCCACCCCTCTGGACAGGACATGC 1991
DB |||||||

RESULT 11

AAS40882

ID AAS40882 standard; cDNA; 1118 BP.

XX

AC AAS40882;

XX

DT 17-DEC-2001 (first entry)

XX

DE cDNA encoding novel human enzyme polypeptide #98.

XX

KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.

XX Homo sapiens.
OS WO200155301-A2.
XX
PN 02-AUG-2001.
PD
PF 17-JAN-2001; 2001WO-US01239.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220964.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 14-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.


```
|||||
484 CGAACTGTGGAGCTGGGCTGACAGCAACCATGAGTTGTCACCCAGTCTTCGGGAC 543
1489 AGTGCCAAACAGTTTGGAGGAAACACCTACTCTTGGCGCGGACCGTGAGACACTGCTG 1548
544 AGTGCCAAACAGTTTGGAGGAAACACCTACTCTTGGCGCGGACCGTGAGACACTGCTG 603
1549 CTCGCGTTTGGCAAGACCATCATGAGGAGCAGCTGTTACTCAAGGGGTGCGCAACATC 1608
604 CTCGCGTTTGGCAAGACCATCATGAGGAGCAGCTGTTACTCAAGGGGTGCGCAACATC 663
1609 CTCATCAACCTGTATGGCATGACGGCGCTGCTGCGGGCGCAGCGCTTCCATCCGCAAT 1668
664 CTCATCAACCTGTATGGCATGACGGCGCTGCTGCGGGCGCAGCGCTTCCATCCGCAAT 723
1669 GGGCTCCGCAACACACACAGCAGGTTCTTGGGCAACACCTTCTGCGTGAAGCTTAC 1728
724 GGGCTCCGCAACACACAGCAGGTTCTTGGGCAACACCTTCTGCGTGAAGCTTAC 783
1729 TTGCAGAACTCTTTCAGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAACCTAGATGAG 1788
784 TTGCAGAACTCTTTCAGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAACCTAGATGAG 843
1789 CAGATTAGAAAGTGTCCAGCAGATCTCTTGAAGCAGGACCTATATCTGTGCCACCT 1848
844 CAGATTAGAAAGTGTCCAGCAGATCTCTTGAAGCAGGACCTATATCTGTGCCACCT 903
1849 CTGGACAGGACATGC 1863
904 CTGGACAGGACATGC 918

RESULT 12
AAH08631
ID AH08631 standard; cDNA; 771 BP.
XX AAH08631;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:5466.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
WPI; 2001-318749/34.
XX
PR primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 5466; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
```

```
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 771 BP; 204 A; 187 C; 219 G; 158 T; 3 other;
```

```
Query Match 37.8%; Score 704.6; DB 22; Length 771;
Best Local Similarity 98.9%; Pred. No. 7.7e-179;
Matches 718; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
```

```
QY 1 ATGAGCGGCTGGGGCTCTTCTTGGCGACACCGGCTGCGGCTGCGGCTGCGGCGGCTG 60
DB 47 ATGAGCGGCTGGGGCTCTTCTTGGCGACACCGGCTGCGGCTGCGGCTGCGGCGGCTG 106
QY 61 GTGGTCTCTACCGCGAACCAGCGGCTGCTGCGGCTGCGGCTGCTGCGGCTGCGGCTG 120
DB 107 GTGGTCTCTACCGCGAACCAGCGGCTGCTGCGGCTGCGGCTGCTGCGGCTGCGGCTG 166
QY 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAAGAGTTTCCCATTTCCAGAAAGTTAGC 180
DB 167 AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAAGAGTTTCCCATTTCCAGAAAGTTAGC 226
QY 181 CAAGATGAACCTAATGAATCAATCAATGTTCTTGGGACCCGCTGAGGCTTCTCACTGAA 240
DB 227 CAAGATGAACCTAATGAATCAATCAATGTTCTTGGGACCCGCTGAGGCTTCTCACTGAA 286
QY 241 GAGGTGACTCCCGGAAATTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 287 GAGGTGACTCCCGGAAATTTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 346
QY 301 TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAAGAGATATGTTGGCTTGGGCTTTC 360
DB 347 TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAAGAGATATGTTGGCTTGGGCTTTC 406
QY 361 TCCAACACCATGTACTCAAGCTAGGAGATCATGAGTGGTGGTGGTGGTGGTGGTGGTGG 420
DB 407 TCCAACACCATGTACTCAAGCTAGGAGATCATGAGTGGTGGTGGTGGTGGTGGTGGTGG 466
QY 421 ACCCTGGCAGCGCAGGCTATTGGCTTCAAGGGATCATCTTGGCTGGCATGAGGAG 480
DB 467 ACCCTGGCAGCGCAGGCTATTGGCTTCAAGGGATCATCTTGGCTGGCATGAGGAG 526
QY 481 CAGAAAGCCAAATACCTTGGCTTCAAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 527 CAGAAAGCCAAATACCTTGGCTTCAAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 586
QY 541 ACSGAGCAGCCAGTGGGAGGATGCAAGCTCAATCCGAGCAGGAGGAGGAGGAGGAGGAG 600
DB 587 ACSGAGCAGCCAGTGGGAGGATGCAAGCTCAATCCGAGCAGGAGGAGGAGGAGGAGGAG 646
QY 601 GACAAGAGCAGCTACATCTCAATGGCTTCAAGGCTTGGATTTACTATGAGGAGGAGGAG 660
DB 647 GACAAGAGCAGCTACATCTCAATGGCTTCAAGGCTTGGATTTACTATGAGGAGGAGGAG 706
QY 661 AATATTTTACTGTGTTTGCAAGAGCTGAGGTCTGTTGATGATGATGATGATGATGATGAT 720
```

```
|||||
Db 707 AATATTTTACTGGGTTGCNAAGACTGANGTCG-TGATTCTGATGCATCAAGAC 765
Qy 721 AAAATC 726
Db 766 AAAATC 771

RESULT 13
AAS40879
ID AAS40879 standard; cDNA; 651 BP.
AC AAS40879;
XX
DT 17-DEC-2001 (first entry)
XX
DE cDNA encoding novel human enzyme polypeptide #95.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01239.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0188974.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
```


PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 DR WPI; 1999-430243/36.
 XX
 XX
 XX
 XX
 XX
 PS New isolated human polynucleotides
 PS Claim 1; Page 503; 591pp; English.
 XX
 XX This invention describes novel isolated human polynucleotides obtained
 CC by screening for differential expression in colon cancer, breast cancer
 CC and lung cancer cell lines. The polynucleotides of the invention are
 CC represented in AA98275-X99118 and encode polypeptides of protein
 CC families selected from 4 transmembrane segments integral membrane
 CC proteins, 7 transmembrane receptors, ATPases associated with various
 CC cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of
 CC transcription factors, G-protein alpha subunit, phospholipase or
 CC diacylglycerol binding proteins, protein kinase, protein phosphatase 2C,
 CC signalling proteins and WW/rsp5/WMP domain containing proteins. The
 CC encoded polypeptides also have a functional domain selected from Ank
 CC repeat, basic region plus leucine zipper transcription factors,
 CC bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger
 CC (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease
 CC domain. The polynucleotides encode polypeptides with similarity to known
 CC protein families and are predicted to have similar properties. The novel
 CC polynucleotides can be used to develop products for use as therapeutic
 CC agents and in forensics, genetic analysis, mapping and diagnostic
 CC applications. In particular, the product can be used for the detection
 CC and management of cancers. They can be used for treating e.g. cervical
 CC cancers, melanomas, colorectal adenocarcinomas, Wilms' tumour, sarcomas,
 CC retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic
 CC myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and
 CC myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric
 CC hereditary ectodermal dysplasia, congenital alveolar dysplasia,
 CC epithelial dysplasia of the cervix, fibrous dysplasia of bone, and
 CC mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,
 CC prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of
 CC the skin.
 XX
 SQ Sequence 772 BP; 161 A; 207 C; 186 G; 165 T; 53 other;
 Query Match 26.8%; Score 499.6; DB 20; Length 772;
 Best Local Similarity 91.1%; Pred. No. 8.9e-124;
 Matches 545; Conservative 0; Mismatches 49; Indels 4; Gaps 3;
 QY 1032 GAAATTTGCACTGATGGCTCAGAGGCTTACGTATGAGAGTATGACCTACCTCACAGC 1091
 DB 83 GAAATTTGCACTGATGGCTCAGAGGCTTACGTATGAGAGTATGACCTACCTCACAGN 142
 QY 1092 AGGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGCCATGGTGAAGGT 1151
 DB 143 AGGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGAGCCATGGTGAAGGT 202
 QY 1152 GTTCAGCTCCGAGCGCGCTGGCAGTGTGTCAGTGAGCGCTGCAGATCTTCGGGGGCTT 1211
 DB 203 GTTCAGCTCCGAGCGCGCTGGCAGTGTGTCAGTGAGCGCTGCAGATCTTCGGGGGCTN 262
 QY 1212 GGGTACACAAAGGAGTATCCGTACGAGCGCATACTGCGTGACACCGCGCATCTCTCCAT 1271
 DB 263 GGGTACACAAAGGAGTATCCGTACGAGCGCATACTGCGTGACACCGCGCATCTCTCCAT 322
 QY 1272 CTTTCAGGGACCAATGAGATCTCCGGATGATACATCGCCCTGACGGGTCTGCAGCATGC 1331
 DB 323 CTTTCAGGGACCAATGAGATCTCCGGATGATACATCGCCCTGACGGGTCTGCAGCATGC 382
 QY 1332 CGGCGGATCTGACTACGAGGATCCATGAGCTTAAACAGCCCAAGTGCACAGTCAT 1391
 DB 383 CGGCGGATCTGACTACGAGGATCCATGAGCTTAAACAGCCCAAGTGCACAGTCAT 442
 QY 1392 GGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGACCTGGGGCTGAC 1451
 DB 443 GGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGACCTGGGGCTGAC 502

QY 1452 AGGCAACCATGAGTTGTGCACCCAGCTTCCGGACAGTGCACACAGTTTGGAGGAGAA 1511
 DB 503 AGGCAACCATGAGTTGTGCACCCAGCTTCCGGACAGTGCACACAAATTTGGAGGAGAA 562
 QY 1512 CACCTACTGCTTCCGGC-CGGACCGTGGAGACACTGCTGCTCCGCTTTTGGC-AAGACCATC 1569
 DB 563 CACCTACTGCTTCCGGC-CGGACCGTGGAGACACTGCTGCTGNTNCCNTTTGGCAAGACCATC 622
 QY 1570 ATGGAGGAGCAGCTGG--TACTGAAAGCGGGTGGCCCAACATCCTCATCAACCTGTATGG 1625
 DB 623 ATGGAGGAGCAGCTGG--TACTGAAAGCGGGTGGCCCAACATCCTCATCAACCTGTATGG 680
 RESULT 15
 AAZ15890
 ID AAZ15890 standard; cDNA; 835 BP.
 AC AAZ15890;
 XX
 DT 12-OCT-1999 (first entry)
 XX
 DE Human gene expression product cDNA sequence SEQ ID NO:3359.
 XX
 KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensics; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9938972-A2.
 XX
 PD 05-AUG-1999.
 XX
 PF 28-JAN-1999; 99WO-US01619.
 XX
 PR 03-APR-1998; 98US-0080666.
 PR 28-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 PR 03-APR-1998; 98US-0080515.
 XX
 (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 PI
 PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 DR WPI; 1999-494092/41.
 XX
 PT Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 XX
 PS Claim 1; Page 1608; 2479pp; English.
 XX
 XX The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purposes, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and
 CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between

CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.

XX
SQ Sequence 835 BP; 178 A; 212 C; 216 G; 200 T; 29 other;

```
Query Match          26.3%; Score 490.4; DB 20; Length 835;
Best Local Similarity 97.6%; Pred. No. 2.7e-121;
Matches 560; Conservative 0; Mismatches 8; Indels 6; Gaps 6;

QY 1015 TTGGATTGATTCAGAGAAATTTGCACTGCTGCTCAGAGGCTTACGTCATGGAGAG 1073
DB 58 TTGGATTGATTCAGAGAGAAATTTGCACTGATGGCTCAGAA-GCTTACGTCATGGAGAG 116

QY 1074 TATGACCTACCTCAGAGAGGATGCTGACCAACCTGGCTTTCCGACTGCTCCATCGA 1133
DB 117 TATGACCTACCTCAGAGAGGATGCTGACCAACCTGGCTTTCCGACTGCTCCATCGA 176

QY 1134 GGCAGCCATGGTGAAGGTGTTGAGTCCGAGCGCGCTGCGAGTGTGAGTGAGCGCGCT 1193
DB 177 GGCAGCCATGGTGAAGGTGTTGAGTCCCGA-GCCGCTGGCAGTGTGAGTGAGCGCGCT 235

QY 1194 GCAGATCCTCGGGGGCTTGGGCTTACAAAGGGACTATCCGTACGAGCGCATACTGCGTGA 1253
DB 236 GCAGATCCTCGGGGGCTTGGGCTTACAAAGGGACTATCCGTACGAGCGCATACTGCGT-A 294

QY 1254 CACCCGATCCTCCTCATCTTCGAGGAGCAACATGAGATTCCTCGGATGTACATCGCCCT 1313
DB 295 CACCCGATCCTCCTCATCTTCGAGGAGCAACATGAGATTCCTCGGATGTACATCGCCCT 354

QY 1314 GACGGGTCTGCAGCATGCCGGCGCATCCTGACTACGAGATCCATGAGCTTAAACAGGC 1373
DB 355 GACGGGTCTGCAGCATGCCGGCGCATCCTGACTACGAGATCCATGAGCTTAAACAGGC 414

QY 1374 CAAAGTGAGCACAGTATGATACCGTTGGCCGGAGGCTTCGGGACTCCCTGGGCGGAAC 1433
DB 415 CAAAGTGAGCACAGTATGATACCGTTGGCCGGAGGCTTCGGGACTCCCTGGGCGGAAC 474

QY 1434 TGTGGACCTGGGGCTGACAGGCAACCATGGAGTTGTGACCCCGAGTCTTGGGACAGTGC 1493
DB 475 TGTGGACCTGGGGCTGACAGGCAACCATGGAGTTGTGACCCCGAGTCTTGGGACAGTGC 534

QY 1494 CAACAAGTTTGAGGAGAACACCTACTGCTTCGCCGGA-CCGTGGAGACAC-TGCTGCTC 1551
DB 535 CAACAAGTTTGAGGAGAACACCTACTGCTTCGCCGGAACCCGTGGAGACACTTNTTGTTC 594

QY 1552 CGCTTTGGCAAGACCATCATGAGGAGCAGCTGG 1585
DB 595 CGCTTTGGCAAGACCATCATGAGGAGCAACACTTG 628
```

Search completed: September 6, 2003, 17:43:06
Job time : 518.782 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:17:56 ; Search time 26 Seconds
(without alignments)
1123.215 Million cell updates/sec

Title: US-09-945-326-2

Perfect score: 3153

Sequence: 1 MSGCGLFLRTAARACRGL.....SQ0ILEKRAYICAHPLDRTC 621

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1360	43.1	656	1	ACDV_MOUSE
2	1356.5	43.0	655	1	ACDV_RAT
3	1345	42.7	655	1	ACDV_BOVIN
4	1308	41.5	655	1	ACDV_HUMAN
5	711	22.5	650	1	Y873_MYCTU
6	691.5	21.9	379	1	ACDB_BACSU
7	652.5	20.7	379	1	ACDA_BACSU
8	630	20.0	389	1	ACDP_MYCTU
9	625.5	19.8	379	1	ACDS_CLOAB
10	609	19.3	389	1	ACDP_MYCLE
11	604.5	19.2	401	1	IVD2_SOLTU
12	604	19.2	432	1	ACDE_HUMAN
13	593.5	18.8	383	1	ACDS_MEGEL
14	587.5	18.6	412	1	IVD1_SOLTU
15	587.5	18.6	423	1	IVD_HUMAN
16	585.5	18.6	424	1	IVD_RAT
17	583	18.5	412	1	ACDS_HUMAN
18	577	18.3	413	1	ACDS_PIG
19	575	18.2	412	1	ACDS_RAT
20	575	18.2	417	1	ACDB_CABEL
21	575	18.2	432	1	ACDB_MOUSE
22	572.5	18.2	432	1	ACDB_MOUSE
23	567	18.0	412	1	ACDS_MOUSE
24	546.5	17.3	409	1	IVD_ARATH
25	546.5	17.3	421	1	ACDM_HUMAN
26	542.5	17.2	421	1	ACDM_MOUSE
27	529.5	16.8	421	1	ACDM_PIG
28	525	16.7	419	1	ACDM_DROME
29	523	16.6	421	1	ACDM_RAT
30	507	16.1	413	1	ACDB_MOUSE
31	479	15.2	415	1	ACDB_HUMAN
32	444	14.1	430	1	ACDL_RAT
33	441	14.0	430	1	ACDL_PIG

RESULT 1

ID	ACDV_MOUSE	STANDARD;	PRT;	656 AA.
AC	P50544; Q35289; O55133;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Acyl-CoA dehydrogenase, very-long-chain specific, mitochondrial precursor (EC 1.3.99.-) (VLCAD) (MVLCD).			
GN	ACADVL OR VLCAD.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/SVJ; TISSUE=Blood;			
RA	Andresen B.S., Lund H., Bross P., Gregersen N.;			
RT	"Mouse very-long-chain acyl-CoA dehydrogenase (VLCAD) gene.";			
RL	Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c; TISSUE=Heart;			
RA	Andresen B., Lund H., Bross P., Corydon M., Gregersen N.;			
RT	"Cloning and characterization of mouse very-long-chain acyl-CoA dehydrogenase.";			
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 84-656 FROM N.A.			
RC	STRAIN=ICR; TISSUE=Liver;			
RA	MEDLINE=98345418; PubMed=9680378;			
RT	Cox K.B., Johnson K.R., Wood P.A.;			
RL	"Chromosomal locations of the mouse fatty acid oxidation genes Cpt1a, Cpt1b, Cpt2, Acadvl, and metabolically related Crat gene.";			
RN	[4]			
RP	SEQUENCE OF 339-656 FROM N.A.			
RA	Rao G., Krimer D., Krasikov T., Austin C., Skoultschi A.I.;			
RT	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.			
CC	FUNCTION: ACTIVE TOWARD ESTERS OF LONG-CHAIN AND VERY-LONG CHAIN FATTY ACIDS SUCH AS PALMITOYL-COA AND STEAROYL-COA.			
CC	FUNCTION: CATALYTIC ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced ETF			
CC	COFACTOR: FAD.			
CC	PATHWAY: Mitochondrial fatty acid beta-oxidation system; first step.			
CC	SUBUNIT: Homodimer (By similarity).			
CC	SUBCELLULAR LOCATION: Mitochondrial inner membrane.			
CC	MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN TISSUES.			
CC	SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its			

P28330 homo sapien
P31174 mus musculu
Q89b20 proteus sp.
P34275 caenorhabdi
P31571 escherichia
Q82912 salmonella
Q82r22 salmonella
Q80772 caenorhabdi
Q92947 homo sapien
P81140 sus scrofa
Q82937 salmonella
Q8zby6 yersinia pe

34 440.5 14.0 430 1 ACCL_HUMAN
35 437 13.9 430 1 ACCL_MOUSE
36 391 12.4 380 1 CAIA_PROSL
37 390.5 12.4 408 1 IVD_CAEEL
38 390 12.4 380 1 CAIA_ECOLI
39 386 12.2 380 1 CAIA_SALTI
40 386 12.2 380 1 CAIA_SALTY
41 385 12.2 409 1 GCDH_CAEEL
42 383.5 12.2 438 1 GCDH_HUMAN
43 361 11.4 408 1 GCDH_PIG
44 357.5 11.3 814 1 FADE_SALTI
45 357.5 11.3 815 1 FADE_YERPE

ALIGNMENTS

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL; Y11770; CAA72435.1; --
 DR EMBL; 271189; CRA94919.2; --
 DR EMBL; AF017176; AAC31642.1; --
 DR EMBL; U41497; AAB85185.1; --
 DR HSSP; Q06319; IBCU.
 DR SWISS-2DPAGE; P50544; MOUSE.
 DR MGD; MG1:895149; Acadv1.
 DR InterPro; IPR006089; Acyl-CoA.dh.
 DR InterPro; IPR006090; Acyl-CoA.dh.C.
 DR InterPro; IPR006091; Acyl-CoA.dh.M.
 DR InterPro; IPR006092; Acyl-CoA.dh.N.
 DR Pfam; PF00441; Acyl-CoA.dh; 1.
 DR Pfam; PF02770; Acyl-CoA.dh.M; 1.
 DR Pfam; PF02771; Acyl-CoA.dh.N; 1.
 DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
 DR Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
 KW Mitochondrion; Transit peptide.
 FT TRANSIT 1 41 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 42 656 ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN
 FT SPECIFIC.
 FT DOMAIN 42 483 CATALYTIC.
 FT CONFLICT 339 341 NNG -> GTR (IN REF. 4).
 FT CONFLICT 423 423 C -> W (IN REF. 4).
 FT CONFLICT 427 427 A -> G (IN REF. 4).
 FT CONFLICT 441 441 M -> I (IN REF. 4).
 FT CONFLICT 507 507 G -> A (IN REF. 4).
 FT CONFLICT 532 532 R -> P (IN REF. 4).
 FT CONFLICT 532 532 Q -> K (IN REF. 4).
 FT CONFLICT 567 567 AD -> GG (IN REF. 4).
 FT CONFLICT 570 571 A -> P (IN REF. 4).
 FT CONFLICT 573 573 G -> A (IN REF. 4).
 FT CONFLICT 593 593 T -> A (IN REF. 4).
 FT CONFLICT 596 596 A -> P (IN REF. 4).
 FT CONFLICT 612 612 H -> Q (IN REF. 3).
 FT CONFLICT 628 628 H -> Q (IN REF. 3).
 FT SEQUENCE 656 AA; 70875 MW; A0110CA5C6CF4F89 CRC64;

Query Match 43.1%; Score 1360; DB 1; Length 656;
 Best Local Similarity 48.2%; Pred. No. 3.7e-82;
 Matches 287; Conservative 106; Mismatches 184; Indels 18; Gaps 9;
 35 PVRA----PAKELFGKIKKKEVFPPEV-SODELNEINQFLGPEKFFTEVDSRKIDQ 89
 66 PARAEKSFVQWFKGQLTIDQVFPYPSVLSQEQFLKELGVPVAFEEVNDPAKND 125
 90 EGKIPDETLEKLSGLFLGQVPEYGGFGNTMYSRIGEISM-DGSITVTAAHQAI 148
 126 LEKVEDDTLQGLKELGAFGLGFLGSLGSLNTQYARLAEIVGMHDLGVSVTLAGHOSI 185
 149 GLKGIILATEQKAKYLPKLAGSEHIAFCLTEPASGSDAASIRSRATLSEDKHYILN 208
 186 GFKGILLYGTAQREKYLPRVSGAALAAFCITEPSSGSDVASIRSSAIPSCGKIYILN 245
 209 GSKWITNGLANIFTVFAKTEVD-SDGSVKDKITAFIVDFGVGTNGKPEDKLIGRG 267
 246 GSKIWLSNGLADIFTVFAKTPKDAATCAVKEKITAFVVERSGVTHGLPEKMGIIKA 305
 268 SNTCEVHFENKIPVENIIGEVDGPKFVAMNINLSGRFSMGVAGGLKRLLEMTAEYAC 327
 306 SNTSEYFDGVKVPSENVLGEVDGPKFVAVNINLNGRFGMAATLAGTMSKSLIAKAVDHAT 365
 328 TRQFNKRLSEGLICEKALMAQKAYVNESMTYTLTAGMDQPGPDCSIEAAMVKVFS 387
 366 NRTQFGDKLHNGVIGTEKRLARMAILOVYTESMAYMLSANWDQ-GFKDFQIEAAISKIFCS 424
 388 EAAWOCVSEALQILGLGTRDYPYERILRDLTRILLIFEGTNEILRMVYALTGLQHGRI 447
 425 EAAWKADECIQIMGGMGFMKPGVERVLDIRIFRIFEGANDILRLFVALQGMCKGKE 484

QY 448 LTRTRHELKQ--AKVSTVMDTVGRRRLRSLGRTVDLGLTGNHGVVHPSLADSANKFEENT 505
 DB 485 LTGLGNALKNPFNGVGLLMGEAGKQLRRRTGIGSGLSLS---GIVHPELSRSGELAVQAL 541
 QY 506 YCFGRVTETLLLRFGKTIMEEQVLVKRVANILNLVGMTAVLVSRAISRIRGLRHHDEHV 565
 DB 542 DQFATVVEAKLVKHKHGIIVNEQFLQLRLADGADLYAMVVVLVSRAISRIRGLRHHDEHV 601
 QY 566 LLANTFCVEA---YLQNLFSLSQLDKYAPENLDEQIKKVSQOILEKRAYICAHPL 617
 DB 602 MLCDSWCIEAATRIENNASLQSSPQH--QELFNFRFSISKAMVENGGLVTGNPL 654
 RESULT 2
 ID ACV_RAT STANDARD; PRT; 655 AA.
 AC AC P45953;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acyl-CoA dehydrogenase, very-long-chain specific, mitochondrial
 DE precursor (EC 1.3.99.-) (VLCAD).
 GN ACADVL OR VLCAD.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=WiStar; TISSUE=Liver;
 RX MEDLINE=94308174; PubMed=8034667;
 RA Aoyama T., Ueno I., Kamijo T., Hashimoto T.;
 RT "Rat very-long-chain acyl-CoA dehydrogenase, a novel mitochondrial
 RT acyl-CoA dehydrogenase gene product, is a rate-limiting enzyme in
 RT long-chain fatty acid beta-oxidation system. cDNA and deduced amino
 RT acid sequence and distinct specificities of the cDNA-expressed
 RT protein." *Biol. Chem.* 269:19088-19094(1994).
 RL J. Biol. Chem. 269:19088-19094(1994).
 CC -!- FUNCTION: ACTIVE TOWARD ESTERS OF LONG-CHAIN AND VERY-LONG CHAIN
 CC FATTY ACIDS SUCH AS PALMITOYL-COA AND STEAROYL-COA.
 CC -!- CATALYTIC ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced
 CC ETF.
 CC -!- COFACTOR: FAD.
 CC -!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
 CC step.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES
 CC OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
 CC TISSUES.
 CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; D30647; BAA06331.1; --
 DR PIR; A54872; A54872.
 DR HSSP; Q06319; IBCU.
 DR InterPro; IPR006089; Acyl-CoA.dh.
 DR InterPro; IPR006090; Acyl-CoA.dh.C.
 DR InterPro; IPR006091; Acyl-CoA.dh.M.
 DR InterPro; IPR006092; Acyl-CoA.dh.N.
 DR Pfam; PF00441; Acyl-CoA.dh; 1.
 DR Pfam; PF02770; Acyl-CoA.dh.M; 1.
 DR Pfam; PF02771; Acyl-CoA.dh.N; 1.
 DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE; PS00073; ACYL_COA_DH_2; 1.

KW Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
KW Mitochondrion; Transit peptide.
FT TRANSIT 1 40
FT CHAIN 41 655
FT DOMAIN 41 482
FT CATALYTIC SPECIFIC.
SQ SEQUENCE 655 AA; 70749 MW; E808DEB0E4595D7 CRC64;

Query Match 43.0%; Score 1356.5; DB 1; Length 655;
Best Local Similarity 46.4%; Pred. No. 6.3e-82;
Matches 295; Conservative 108; Mismatches 192; Indels 41; Gaps 12;

QY 9 RTTAAACRGLVSTANRRLLR-----TSPVRA-----FAKELGKIKKKEV 53
DB 32 RPTSAQR-----LYASEATQAVLEKPELTSSDASTREKPARAESKSFVGMFKGQITDQV 87
QY 54 PFPPEVSODELINE-----INQFLGPVKEFTTEEDVSRKIDEGKIPDETLEKLSGLGFG 108
DB 88 FYPFSV-----LNEGQTOFLKELGVPAFFVEEVNDPAKNDLSLEKVEEDTLQGLKELGAFG 143
QY 109 LQVPEYGGGLGFSNTMYSLRGEIISM-DGSITVTTLAAHQAATGLKGIILAGTPEEKAKYLP 167
DB 144 LQVPSGLGLGSLNTQVARLAEIVGMHDLGVSVTLGAHQSGIFGKILLYGTAKOKEKYL 203
QY 168 KLASGEHTAACLTEPASGSDAASIRSRATLSEDKHYILNGSKWVTINGLANIFTVFA 227
DB 204 RVASGOALAFCLTEPSSGSDVASIRSAVPSPCGKYTLNGSKIWIISNGGLADIFTVFA 263
QY 228 KTEVVD-SDGSVKDKITAFIVERDFGGVTNCKPKDKIGRSCNTCEVHFNTKIPVNIL 286
DB 264 KTPDKAATGAVKEKITAIVVERSEFGVTHGLPEKMKIGIKASNTSEVYFDGKVPADNVL 323
QY 287 GEVGDGFKVAMNINLNGRFSNGSVYAGLLKRLIEMTAIEYACTRQFNKRLSEFGLIOEKF 346
DB 324 GEVGDGFKVAMNINLNGRFSNGSVYAGLLKRLIEMTAIEYACTRQFNKRLSEFGLIOEKL 383
QY 347 ALMAKAVVEMSTYLTAGMLDQGFDCSIEAMKVFSEAAQVCSALQILGLGY 406
DB 384 ARMAILQVTEVMYMLISANDQ-GFKDFQIEAIAISIFGSEAAKVTDECIQIMGMGF 442
QY 407 TRDYPYRILDRTRILLIFECTNEILMYIALTGLQAHGRILITRIHELKQ--AKVSTVM 464
DB 443 MKPGVERVLDIRIPFESTNDILRLFVALQCMKDKELTGLGNALKNPLGNVGLLI 502
QY 465 DTGVRRLRDSLGRTVDLGLTNGVGVHPSLADSANKFEENTYCFGRVETLLRFRGKTIM 524
DB 503 GEASKQURRTGTGSGLSL--GIVHPELSRSGELAVQALEQFATVWEAKLMKHKKGV 559
QY 525 EEQLVLRVANILINLYGMTAVLSRASRSTRIGLRNHDHDEVLANTFCVEA---YLQNL 581
DB 560 NEQFLQRLADGALDLYAMVVLVSRASRSLSEGYPTAQHEKMLCDSWCIEAATRIENMA 619
QY 582 SLSLODKAPENLDEQIKKYSQOILEKRAYICAHPL 617
DB 620 SLSQSNPQ--QOELFRNFRSISKAMVENGGLVTSNPL 653

RESULT 3
ACDV_BOVIN STANDARD; PRT; 655 AA.
AC P48818;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, very-long-chain specific, mitochondrial precursor (EC 1.3.99.-) (VLCD).
GN ACADVL OR VLCD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Zhang X., Liu W., Zhu H., Sun X.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVE TOWARD ESTERS OF LONG-CHAIN AND VERY-LONG CHAIN
CC FATTY ACIDS SUCH AS PALMITOYL-COA AND STEAROYL-COA.
CC -!- CATALYTIC ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced
CC ETF.
CC -!- COFACTOR: FAD.
CC -!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
CC step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES
CC OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
CC TISSUES.
CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U30817; AAA74051.1; -;
DR HSP; Q06319; IBCU.
DR InterPro: IPR006089; Acyl-CoA_dh.
DR InterPro: IPR006090; Acyl-CoA_dh_C.
DR InterPro: IPR006091; Acyl-CoA_dh_M.
DR InterPro: IPR006092; Acyl-CoA_dh_N.
DR Pfam: PF00441; Acyl-CoA_dh; 1.
DR Pfam: PF02770; Acyl-CoA_dh_M; 1.
DR Pfam: PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE: PS00072; ACYL_COA_DH_1; 1.
DR PROSITE: PS00073; ACYL_COA_DH_2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
KW Mitochondrion; Transit peptide.
FT TRANSIT 1 40
FT CHAIN 41 655
FT DOMAIN 41 482
FT CATALYTIC SPECIFIC.
SQ SEQUENCE 655 AA; 70520 MW; F9DC06285023CFC0 CRC64;

Query Match 42.7%; Score 1345; DB 1; Length 655;
Best Local Similarity 47.3%; Pred. No. 3.6e-81;
Matches 285; Conservative 110; Mismatches 194; Indels 14; Gaps 9;

QY 23 STANRRLLRTSPVRAFAKELFLGKIKKEVFPPEV-SQDELNEINQFLGPVKEFTPEE 81
DB 57 SEASTREKRANSVSKSFVGTGKQGLTDDQVFPVPSVLNEDQTFKLKELGVPVTFEEV 116
QY 82 VDSRKIDCGKIPDETLEKLSGLGFLGVPEYGGGLGFSNTMYSLRGEIISM-DGSITV 140
DB 117 NDAAKNDMLERVEETTMQGLKELGALGVLPNLPNGLCQYARLVEIVGMVLDGVGI 176
QY 141 TLAAHQAIKGLIILAGTPEEKAKYLPKLASGEHTAACLTEPASGSDAASIRSRATLSE 200
DB 177 VLGAHQSGIKGLLFGTKAQKELYPKLASGETTAACFLTEPSSGSDAASIRSAVSPSP 236
QY 201 DKKHYILNGSKWVTINGGLANIFTVFAKTEVVD-SDGSVKDKITAFIVERDFGGVTNKP 259
DB 237 CKKYVTLNGSKWISNGGLADIFTVFAKTPVTDATGAVKEKITAIVVERSEFGVTHGPP 296
QY 260 EDKLGIRGNTCEVHFNTKIPVENILGVGGFKVAMNINLNGSRFSNGSVVAGLLKRI 319
DB 297 EKKMGIKASNTAEVYFDGVRVPAENVLGVGGFKVAMHILNNGRFGMAALAGATMKGII 356
QY 320 EMTEVACTRKQFNKRLSEFGLIOEKFALMAQKAVYMESMTYLTAGMLDQGFDCSIEA 379
DB 357 AKAVDHAANRTQFGEKIHFGLIQLKELARMAMQLQVVTESMAYMVSANMDQ-GSTDQFIEA 415


```

RA RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.: "
RT Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RN Nature 393:537-544(1998).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=CDC 1551 / Oshkosh;
RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bisbal W.;
RT "whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains";
RL submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC -I- COFACTOR: FAD (BY SIMILARITY).
CC -----
CC -I- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outpost at
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL022004; CAA16769.1; -
DR EMBL; AE006977; AAU45138.1; -
DR PIR; A70817; A70817.
DR HSP; Q06319; IBUC.
DR TIGR; MT0896; -.
DR TubercuList; RV0873; -.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh_1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; FALSE-NEG.
DR PROSITE; PS00073; ACYL_COA_DH_2; FALSE-NEG.
DR Hypothetical protein; Oxidoreductase; Flavoprotein; FAD;
KW Complete proteome. 464
FT DOMAIN 443
SQ SEQUENCE 650 AA; 70744 MW; FCF934A3F01C0D1 CRC64;
Query Match 22.5%; Score 711; DB 1; Length 650;
Best Local Similarity 31.4%; Pred. No. 2.e-39;
Matches 194; Conservative 113; Mismatches 223; Indels 88; Gaps 16;

QY 38 AFAKELEFLGKTKKEVFPFFPVODELNENQLPGVEKPTEEVDSRKIDDEGKIIPDET 97
DB :|||||:: ::|||:: |||:: ::|||:: |||:: |||:: |||:: |||:: |||:: |||::
DB SFAKELFLGRPLGLIIHPFKPSDAEEARTAEFLVKLRFL-DTVDGSVIERAAQIPDEY 86

QY 98 LEKLKSLLGLFGLOVPPEEYGGSGFNMTWYSR-LGEIISMDGSITVTLAHQHAIGL-KGIIL 155
DB ::|||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||::
DB VKGLAEICFCGLKPISEYGGNLNMQAVNYRVLMVTTVVHSSGALLSAHQSIGVPEPLK 146

QY 156 AGTEEQAKAYLPKLIASGHIAAFCLITEPASGDRAASTRSATLSDEKKHYILNGSKWIT 215
DB |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||:: |||::
DB 147 AGTAEQKRRLFPCRAAGA-ISAFLLITEPDVGSDPFARMASATAPIDDGOAYELEGVKLWT 205

QY 216 NGGLANIFTVFaktevvdsgsvdkdktafivverdfgvgtngkpedklgrgsntcevhf 275
DB ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::
DB NGVVDLLVWAR--VPRSEG-RGGISAFVVEADSPGITVERRNKFMGLRIENGIVTRL 262

QY 276 ENTKTPVENILGEVDFGFKAAMNIINSGRFSMGVSVAAGLKRLIMETIACTRKQPNKR 335
DB ::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::
DB 263 HRVRVPKNLLIGRGDGLKIALTTNACRLSLPALATGVAQAALKIAREWSVERVQWGPK 322

QY 336 LSEFGLIOEKPAALMAQAAYVMESMTHLTAGMLDPGPDCSTEAAMKVFSSEAAMOCVS 395
DB ::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::

```

	QY	335 RLSFGLQLQRFPALMAOKAYWESWTLYTTAGMLDQPFGPDCSTEAAMVKVFSEAAQCVCV	394	
	Dd	: : :		
	Db	272 PIAANQGISFKLADMATRAEAAHRLVHAAD-LHNRLG-NCGRKASNAQOFASDAAVKA	329	
	:	: : : :		
	QY	335 SEALQLGGIGYTROYPERILRDRTRILLIFEGFTNETLRMYIA	437	
	Dd	: : :		
	Db	330 LDAVOIYGSGYMKDPVERLLRRDAKVTOYEGTNEIQRIILIIS	372	
	:	: : :		
	 RESULT 7 			
ID	ACDA_BACSU	STANDARD;	PRT; 379 AA.	
AC	P45867:			
DT	01-NOV-1995 (Rel. 32, Created)			
OT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	AcyI-CoA dehydrogenase (EC 1.3.99.-).			
GN	ACDA OR ACD.			
OS	Bacillus subtilis.			
OX	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OC	NCBII_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=98015417; PubMed=9353933;			
RR	Presecan E., Moszer I., Boursier L., Cruz Ramos H.C., De La Fuente V.,			
RR	Hullo M.-F., LeLONG C., Schleich S., Sekowska A., Song B.H.,			
RR	Villani G., Kunst F., Danchin A., Glaser P.;			
RT	"The bacillus subtilis genome from gerBC (31 degrees) to lncR (334			
RT	degrees).";			
RL	Microbiology 143:3313-3328(1997).			

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; Z49782; CAA89868.1; -
 EMBL; Z99123; CAB15745.1; -
 PIR; S5421; S53421.
 HSSP; Q06319; 1BUC.
 DR Subtilist; BG11239; acda.
 DR InterPro; IPR006089; Acyl-CoA_dh.
 DR InterPro; IPR006090; Acyl-CoA_dh.C.
 DR InterPro; IPR006091; Acyl-CoA_dh.M.
 DR InterPro; IPR006092; Acyl-CoA_dh.N.
 DR Pfam; PF00441; Acyl-CoA_dh; 1.
 DR Pfam; PF02770; Acyl-CoA_dh.M; 1.
 DR Pfam; PF02771; Acyl-CoA_dh.N; 1.
 DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
 KW Oxidoreductase; Flavoprotein; FAD; Complete proteome.
 SQ SEQUENCE 379 AA; 41446 MW; 4D09861D59718EF9 CRC64;

Query Match 20.7%; Score 652.5; DB 1; Length 379;
 Best Local Similarity 42.1%; Pred. No. 7.4e-36;
 Matches 155; Conservative 60; Mismatches 140; Indels 13; Gaps 7;

74 VEKFTTEV--DSRKIDGKIPDETLEKLSGLFGLQVPEEYGGGLFSNTMY-SRLGE 130
 16 VRDFAHEVAPTAAERDEQERDFRELFREMANLGTGIPWDEYGGIGSDYLAVIAVEE 75
 131 IISMDGSIIVTAAHOATGLKGIILAGTEQAKYLPKLSAGEHIAAFCLTEPAGSDAA 190
 76 LSKVCASGTVLSAHLSICSWPLFAFGTEQKTEVLTQLGKIGAFALTEAGSDAG 135
 191 SIRSRLATSEDKKKHYLNGSKVWITNGLANIFTVFATKTEVVDSDGSKDK-ITAFIVER 249
 136 SMKTTAERIGD--DYVLNGSKVFTNGGVADIYVFAVT-----DPEKKKKGVTAIVFK 188
 250 DFGGVNGKPKDLGRSNTCEVHFENTKIPVENILGEVGDGFKVAMNINLSGRFSMG 309
 189 DEGFFTGKKEKGLIRSPTEIMFEDCVVPASKRLGEEGEGFKIAMKTLDDGGRNGIA 248
 310 VVAGLLKRLIEMTAETACTKRFNKRLESEGLIQKFAALMAQKAVVMESWYLYFAGLDQ 369
 249 QAVGQAQALDAALQAKERKQFGKSIARQOQIAFAFLADMATMIEASRLITY-QAAWLES 307
 370 PGFPDCSIEAMVKVFSSEAAQCVCSEALQILGLGYTRDYPYERILRTRILLIFEGTN 429
 308 SGLP-YGKASAMSKLMAGDTAMKVTEAVQIFGGYTKDYPPVRYMRDAKITQIYEGTQ 366
 430 EILRMVIA 437
 367 EIQLVIS 374

RESULT 8
 ACDP_MYCTU STANDARD; PRT; 389 AA.
 ID ACDP_MYCTU
 AC P56879;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable acyl-CoA dehydrogenase fadE25 (EC 1.3.99.-).
 GN FADE25 OR ACD OR RV3274C OR MT3374 OR MTCV71.14C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares R., Squares R.,
 Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RT Nature 393:537-544 (1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CDC 1551 / Oshkosh;
 RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
 CC ETF
 CC -1- COFACTOR: FAD (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ACYL-CoA DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

EMBL; Z92771; CAB07077.1; -
 EMBL; AE007147; AAK47715.1; -
 PIR; C70979; C70979.
 HSSP; Q06319; 1BUC.
 DR TIGR; MT3374; -
 DR Tuberculist; RV3274C; -
 DR InterPro; IPR006089; Acyl-CoA_dh.
 DR InterPro; IPR006090; Acyl-CoA_dh.C.
 DR InterPro; IPR006091; Acyl-CoA_dh.M.
 DR InterPro; IPR006092; Acyl-CoA_dh.N.
 DR Pfam; PF00441; Acyl-CoA_dh; 1.
 DR Pfam; PF02770; Acyl-CoA_dh.M; 1.
 DR Pfam; PF02771; Acyl-CoA_dh.N; 1.
 DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
 KW Hypothetical protein; Oxidoreductase; Flavoprotein; FAD;
 KW Complete proteome.
 SQ SEQUENCE 389 AA; 41723 MW; 2E77D0F28A9C8E08 CRC64;

Query Match 20.0%; Score 630; DB 1; Length 389;
 Best Local Similarity 38.0%; Pred. No. 2.3e-34;
 Matches 147; Conservative 77; Mismatches 149; Indels 14; Gaps 9;

QY 52 EVFPPEVSQDELNEINFLGPVEKFFTEEDVSRKIDGKIPDETLEKLSGLFGLQV 111
 DB 11 DLFLKPE-EHDEMRSAIRAL--AEKEIAPH--AAVDEKARFPEALVALNSSGFNAVHI 65
 QY 112 PEYGGGLGFSNTMYS-RLGEIISMDGSIIVTAAHOATGLKGIILAGTEQAKYLPKLA 170
 DB 66 PEYGGQAGDSVATCIVTEVARVDASALIPAVNK-LGTMLILRGSEELKKQVLPALA 124
 QY 171 SEGHIAFCLTEPAGSDAAASIRSRATLSEDKKKHYLNGSKVWITNGLANIFTVFAKTE 230
 DB 125 AEGAMASYALSEREAGSDAASMRTRA--KADGDHILNGAKCWCITNGKSWITVYWA--- 179
 QY 231 VVDSGSKVDKITAFIVERDFGVTNGKPEDKILGRSNTCEVHFENTKIPVENILGEV 290
 DB 180 VTDPRGA-NGISAPWVHKDDGFTVGPKERKLGKISPTTELYENCRIPIGDIIEPG 238


```

QY 291 DGFVAMNINLGRSMGSSVAGLLKRLIEMTAETVACTRKQFNKRLSEFGLIOEKPALMA 350
DB 239 TGFXTALATLDHTRPTIGAQAQVGAQALDAAIYTKDRQFGESISTFAVQFMFLADMA 298
QY 351 QKAYVMSWYLYTAGMLDQGFPCDCTEAAVMKVFSEAAQVCSEALQILGLGIGYTRDY 410
DB 299 MKVEAARLMYSAARAER-GEPLGFIISAASKCFASDVAMEVTTDAVQLFGGAGYTTDF 357
QY 411 PYERILDRTRILLFEETNETLRMYIA 437
DB 358 PVERFMDAKITQIYECTNQLQRVMS 384

RESULT 9
ACDS_CLOAB STANDARD; PRT; 379 AA.
AC P52042;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)
DE (SCAD) (Butyryl-CoA dehydrogenase).
DE BCD OR CAC2711.
GN Clostridium acetobutylicum.
OS Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=96236011; PubMed=8655474;
RA Boynton Z.L., Bennett G.N., Rudolph F.B.;
RT "Cloning, sequencing, and expression of clustered genes encoding
beta-hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and
butyryl-CoA dehydrogenase from Clostridium acetobutylicum ATCC
824.";
RT J. Bacteriol. 183:4823-4838(2001).
RL J. Bacteriol. 178:3015-3024(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- CATALYTIC ACTIVITY: Butanoyl-CoA + ETF = 2-butenoyl-CoA + reduced
ETF.
CC -1- COFACTOR:- FAD.
CC -1- PATHWAY: BUTYRATE/BUTANOL-PRODUCING PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U17110; AAA95968.1;
CC DR EMBL; AE007768; AAK80657.1;
CC DR PIR; F97233; F97233.
CC DR PIR; T47262; T47262.
CC DR HSSP; Q06319; 1BUC.
CC DR InterPro; IPR006089; Acyl-CoA_dh.
CC DR InterPro; IPR006090; Acyl-CoA_dh.c.
CC DR InterPro; IPR006091; Acyl-CoA_dh.M.
CC DR InterPro; IPR006092; Acyl-CoA_dh.N.
CC Pfam; PF00441; Acyl-CoA_dh; 1.

```

```

DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Oxidoreductase; Flavoprotein; FAD; Complete proteome.
SQ SEQUENCE 379 AA; 41386 MW; 4BC50A1991BD9FB2 CRC64;

Query Match 19.8%; Score 625.5; DB 1; Length 379;
Best Local Similarity 39.0%; Pred. No. 4.5e-34;
Matches 144; Conservative 73; Mismatches 137; Indels 15; Gaps 8;

QY 74 VEKFFTEEVD--SRKIDQEGKIPDETLEKLSGLGFLQVPEYGGGLGFSNTMY-SRLGE 130
DB 16 VREFAENEVKPIAAEIDETERFPMENVKMGQYMMGIPFSEYGGGADVLSYIAVBE 75

QY 131 IISMDGSITVTTLAAHQAIGLKGIILAGTTEQAKYLPKLASGEHTAAAFCLTPASGSDAA 190
DB 76 LSKVCGTTGCVILSAHTSLCASLINEHGTTEQKQYLVPLAKGEKICAYGLTEPNAGTDSG 135

QY 191 SIRSRTATLSEDKKHVILNGSKVWITNGLANITFTVFAKTEVVDGSGVKDKITAFIVERD 250
DB 136 AOOTVAVEGD--HYVINGSKIFITNGGVADTFVIFAMT--DRTKGTGK-ISAFTIEKG 189

QY 251 FGGVTNGKPEDKLGIRGSNTCEVHEFNTKIPVENILGEVGDGFKVAMNLTNSGFSMSGV 310
DB 190 FKGFSTGKVEQKLGIRASSTTELVEFDMIVPVENMIGKEGKGFPIAMKTLDGGRIGIAAQ 249

QY 311 VAGLLKRLIEMTAETVACTRKQFNKRLSEFGLIOEKPALMAQKAYVMSNTYLT--AGMLD 368
DB 250 ALGIAEGAFNEARAYMKERKQFGRSLDKF--QGLAMNMDADMDVAIESARVLYVKAAYLK 306

QY 369 QGFPDCSTEEAAWKFVSSEAAQVCSEALQILGLGIGYTRDYPPYRILDRTRILLIFGCT 428
DB 307 QAGLP-YTVDAARAKLHAANVAMDVTTRAVQLFGGYGYTKDYPVPMRMDAKITEIYEGT 365

QY 429 NEILRMYIA 437
DB 366 SEVQKLVIS 374

RESULT 10
ACDP_MYCLE
ID ACDP_MYCLE STANDARD; PRT; 389 AA.
AC P46703;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable acyl-CoA dehydrogenase fadE25 (EC 1.3.99.-).
GN FAD25 OR ACD OR ML0737 OR B1308_F1_34.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96084954; PubMed=7489918;
RA Doukhan L., Predlich M., Nair G., Dussurget O., Mandic-Mulec I.,
RA Cole S.T., Smith D.R., Smir I.;
RT "Genomic organization of the mycobacterial sigma gene cluster.";
RL Gene 165:67-70(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajadream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;

```

```

RT  "Massive gene decay in the leprosy bacillus.";
RL  Nature 409:1007-1011(2001).
CC  -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
CC  ETF.
CC  -1- COFACTOR: FAD (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; U00012; AAA85936.1;
CC  EMBL; AL583919; CAC30246.1;
CC  PIR; B87001; B87001.
CC  HSP; Q06319; 1BUC.
CC  Leprosa; M10737.
CC  DR  InterPro; IPR006089; Acyl-CoA_dh.
CC  DR  InterPro; IPR006090; Acyl-CoA_dh_C.
CC  DR  InterPro; IPR006091; Acyl-CoA_dh_M.
CC  DR  InterPro; IPR006092; Acyl-CoA_dh_N.
CC  DR  Pfam; PF00441; Acyl-CoA_dh; 1.
CC  DR  Pfam; PF02770; Acyl-CoA_dh_M; 1.
CC  DR  Pfam; PF02771; Acyl-CoA_dh_N; 1.
CC  DR  PROSITE; PS00072; ACYL_COA_DH_1; 1.
CC  DR  PROSITE; PS00073; ACYL_COA_DH_2; 1.
CC  KW  Hypothetical protein; Oxidoreductase; Flavoprotein; FAD;
CC  Complete proteome.
CC  SQ  SEQUENCE 389 AA; 41719 MW; EFC80CDB3ED884B1 CRC64;

Query Match 19.3%; Score 609; DB 1; Length 389;
Best Local Similarity 36.8%; Pred. No. 5.6e-33;
Matches 143; Conservative 76; Mismatches 152; Indels 18; Gaps 8;

QY 52 EVFFPPEVSQDELNEINQFLGPKFTEEV--DSRKIDQEGKIPDETLEKLSGLFGL 109
DB 11 DLFKLPE-----EHNELRATIRALAEKIAHAADVQARFPEEALAAALNASGFNAI 63

QY 110 QVPEYGGGLGSNTMWS--RLGEIISMDGSIIVTAAHQAGLIGLLAGTEEQAKYLPK 168
DB 64 HYPEYGGOGADSVAACTIVEVARKVDASALIPAVNK-LGTMLGILRGSELKKQVLPS 122

QY 169 LASGEHIAAFCLTEPAGSDAASIRSRATLSDEKHHYILNGSKVITNGGLANIFTVPK 228
DB 123 LAAGAMASYALSEREAGSDAASMTTRA--RADGDDWILNGFKWITNGKSTWTVMA- 179

QY 229 TEVDSGSKVKITAFIVERDFGVTKGPKEDKLGIRGSKNTCEVHFENTKIPVENILGE 288
DB 180 --VTPDKGA-NGISAFIVHKDDGFSIGPKKLGKSGPTTELYFDKCRIPGDRIIGE 236

QY 289 VGDGFKVAMNINLGRFSMGVSVAGLLKRLIEMTAEXACTRKQFNKRLSEFGLIQEKPAL 348
DB 237 PGTGKTALATLDHTRPTGQAQVAGALDAALIVYTKDRKQFGESISTQSIQFMAD 296

QY 349 MAQKAYVMSMTYLTAGMLDQFPDPCSEIAAMVVFSEAAQCVSALQILGLGYTR 408
DB 297 MAMKVEARLIVYAAAAARER-GEPLDGFISAASKCFASDIAMEVTTDAVQLFGAGYTS 355

QY 409 DYPYERILDRTRILLIFEGTNEILRYTA 437
DB 356 DFPVFERMRDAKITQIYEGTNOIQRVMS 384

RESULT 11
ID IVD2_SOLTU
AC Q9FS87;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Isovaleryl-CoA dehydrogenase 2, mitochondrial precursor (EC 1.3.99.10)
DE (IVD 2) (Fragment).
GN IVD2
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
SEQUENCE FROM N.A., SEQUENCE OF 18-44, AND SUBCELLULAR LOCATION.
RP STRAIN=cv. Bintje; TISSUE= tuber;
RX MEDLINE=21153238; PubMed=11331285;
RA Faltre-Nitschke S.E., Couee I., Vermel M., Grienemberger J.M.,
RA Gualberto J.M.;
RT "Purification, characterization and cloning of isovaleryl-CoA
RT dehydrogenase from higher plant mitochondria.";
RL Eur. J. Biochem. 268:1332-1339(2001).
CC -1- CATALYTIC ACTIVITY: 3-methylbutanoyl-CoA + ETF = 3-methylbut-2-
CC enoyl-CoA + reduced ETF.
CC -1- COFACTOR: FAD.
CC -1- PATHWAY: Leucine catabolism.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ278988; CAC08234.1;
CC HSP; P26440; 1IVH.
CC InterPro; IPR006089; Acyl-CoA_dh.
CC InterPro; IPR006090; Acyl-CoA_dh_C.
CC InterPro; IPR006091; Acyl-CoA_dh_M.
CC InterPro; IPR006092; Acyl-CoA_dh_N.
CC Pfam; PF00441; Acyl-CoA_dh; 1.
CC Pfam; PF02770; Acyl-CoA_dh_M; 1.
CC Pfam; PF02771; Acyl-CoA_dh_N; 1.
CC PROSITE; PS00072; ACYL_COA_DH_1; 1.
CC PROSITE; PS00073; ACYL_COA_DH_2; 1.
CC KW  Oxidoreductase; Flavoprotein; FAD; Mitochondrion; Transit peptide.
CC FT  NON-TER 1 17 MITOCHONDRION.
CC FT  TRANSIT <1 17 ISOVALERYL-COA DEHYDROGENASE 2.
CC FT  CHAIN 18 401
CC FT  ACT_SITE 262 262 BY SIMILARITY.
CC SQ  SEQUENCE 401 AA; 43970 MW; 789FF67CA3B4995D CRC64;

Query Match 19.2%; Score 604.5; DB 1; Length 401;
Best Local Similarity 36.8%; Pred. No. 1.2e-32;
Matches 151; Conservative 60; Mismatches 184; Indels 35; Gaps 9;

QY 34 PPVAFAKELFLGRIKKKEVFPPEVSQDELNEINQFLGPKFTEEV--DSRKIDQEG 91
DB 12 PQFAFSTSLFDDTQK-----QKESVAQFAQENIAPHAEKIDRTN 53

QY 92 KIP-DETLEKUK-SIGLFLGLOVPEYGGGLGFSNTMWS--RLGEIISMDGSIIVTAAHQAI 148
DB 54 YFPQDVNLKLMGDFNLLGITVPEYGGGLGYLYHCTAMEEISRASGSLSYGAHTNL 113

QY 149 GLKGTILLAGTEEQAKYLPLKASGEHIAAFCLTEPAGSDAASIRSRATLSDEKHHYILN 208
DB 114 CINQLVRNGTHEQKQKYLPLKILISGEHVALAMSEPDAGSDVVMKCAADRVG--GYVLN 171

QY 209 GSKVITNGGLANIFTVFAKTEVDSGSKVKDKITAFIVERDFGVTKGPKEDKLGIRGS 268
DB 172 GNKMWCTNGPTAQTLVVYAKTDVTTASSKG----ITAFIERKMTGFTFQAQKLDKLMGRGS 227

QY 269 NTCVHFHENTKIPVENILGEVGDGFKVAMNINLGRFSMGVSVAGLLKRLIEMTAEXACT 328
DB 114 CINQLVRNGTHEQKQKYLPLKILISGEHVALAMSEPDAGSDVVMKCAADRVG--GYVLN 171

```

Db 228 DTCELVFENCPEENVLQGVGVVYVLMGLDLRLVLAGSPVGIQACLDVLPYVKQ 287
 QY 329 RKQFNKRLSEFGLOEKALMAQKAVVMESMTYLTAGMLDQP--GFPDCSIEAAMVKVFS 386
 Db 288 REQFGRIPEFQVQKVDMTYMQSSRSYLYSVARECDSTINTKDC----AGVILSA 343
 QY 387 SEAMQCVSEALQIQLGGVYTRDYPYERILRDLTRILLIFEGTNEILRMVY 436
 Db 344 AERATQVALQAIQCLGGNGVYNEPTGRFLRDAKLVGAGTSEIRRMII 393

RESULT 12
 ACDB_HUMAN
 ID ACDB_HUMAN STANDARD; PRT; 432 AA.
 AC P45954; Q96CX7;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Acyl-CoA dehydrogenase, short/branched chain specific, mitochondrial
 DE precursor (EC 1.3.99.-) (SBCAD) (2-methyl branched chain acyl-CoA
 DE dehydrogenase) (2-MEBCAD) (2-methylbutyryl-coenzyme A dehydrogenase)
 DE (2-methylbutyryl-CoA dehydrogenase).
 GN ACADSB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=95213018; PubMed=7698750;
 RA Rozen R., Vockley J., Zhou L., Milos R., Willard J., Fu K.,
 RA Vicanek C., Low-Nang L., Torban E., Fournier B.;
 RT "Isolation and expression of a cDNA encoding the precursor for a
 RT novel member (ACADSB) of the acyl-CoA dehydrogenase gene family";
 RL Genomics 24:280-287(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20489853; PubMed=11013134;
 RA Andresen B.S., Christensen E., Corydon T.J., Bross P., Pilgaard B.,
 RA Wanders R.J.A., Rulter J.P.N., Simonsen H., Winter V., Knudsen I.,
 RA Schroeder L.D., Gregersen N., Skovby F.;
 RT "Isolated 2-methylbutyrylcoinnuria caused by short/branched-chain
 RT acyl-CoA dehydrogenase deficiency: identification of a new enzyme
 RT defect, resolution of its molecular basis, and evidence for distinct
 RL acyl-CoA dehydrogenases in isoleucine and valine metabolism";
 RN Am. J. Hum. Genet. 67:1095-1103(2000).
 RP [3]
 RP SEQUENCE FROM N.A.
 RA Wambutt R., Heubner D., Mewes H.-W., Weil B., Wiemann S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Fahy J., Heiton D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whitting M., Madan A., Young A.C., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP VARIANT 2-METHYLBUTYRYLGLYCINURIA PHE-255.
 RX MEDLINE=20290343; PubMed=10832746;
 RA Gibson K.M., Burlingame T.G., Högema B., Jakobs C., Schutgens R.B.H.,
 RA Millington D., Roe C.R., Roe D.S., Sweetman L., Steiner R.D.,
 RA Linck L., Pohowalla P., Sacks M., Kiss D., Rinaldo P., Vockley J.;
 RT "2-methylbutyryl-coenzyme A dehydrogenase deficiency: a new inborn
 RT error of L-isoleucine metabolism.";
 RL Pediatr. Res. 47:830-833(2000).
 CC -1- FUNCTION: Has greatest activity toward short branched chain acyl-
 CC CoA derivative such as (S)-2-methylbutyryl-CoA, isobutyryl-CoA,
 CC and 2-methylhexanoyl-CoA as well as toward short straight chain
 CC acyl-CoAs such as butyryl-CoA and hexanoyl-CoA. Can use valproyl-
 CC CoA as substrate and may play a role in controlling the metabolic
 CC flux of valproic acid in the development of toxicity of this
 CC agent.
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + ETF -> 2,3-dehydroacyl-CoA + reduced
 CC ETF.
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
 CC step.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- DISEASE: Defects in ACADSB are the cause of 2-
 CC methylbutyrylglucurina [MIM:600301]. This is a defect in
 CC isoleucine catabolism which is characterized by an increase of 2-
 CC methylbutyrylglycine and 2-methylbutyrylcarnitine in blood and
 CC urine.
 CC -1- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U12778; AA74424.1; -
 DR EMBL: AF260678; AAF97921.1; JOINED.
 DR EMBL: AF260678; AAF97921.1; JOINED.
 DR EMBL: AF260669; AAF97921.1; JOINED.
 DR EMBL: AF260670; AAF97921.1; JOINED.
 DR EMBL: AF260671; AAF97921.1; JOINED.
 DR EMBL: AF260672; AAF97921.1; JOINED.
 DR EMBL: AF260673; AAF97921.1; JOINED.
 DR EMBL: AF260674; AAF97921.1; JOINED.
 DR EMBL: AF260675; AAF97921.1; JOINED.
 DR EMBL: AF260676; AAF97921.1; JOINED.
 DR EMBL: AF260677; AAF97921.1; JOINED.
 DR EMBL: AL831821; CAD38535.1; -
 DR EMBL: BC013756; AAH13756.1; -
 DR PIR: A55680; A55680.
 DR HSP: Q06319; 1BUC.
 DR Genew: HGNC:91; ACADSB.
 DR GK: P45954; -
 DR MIM: 600301; -
 DR GO: GO:0005739; C:mitochondrion; TAS.
 DR GO: GO:0003995; F:acyl-CoA dehydrogenase activity; TAS.
 DR GO: GO:0006631; P:fatty acid metabolism; TAS.
 DR InterPro: IPR006089; Acyl-CoA_dh.
 DR InterPro: IPR006090; Acyl-CoA_dh_C.
 DR InterPro: IPR006091; Acyl-CoA_dh_M.
 DR Pfam: PF00441; Acyl-CoA_dh; 1.
 DR Pfam: PF02770; Acyl-CoA_dh_M; 1.
 DR Pfam: PF02771; Acyl-CoA_dh_N; 1.
 DR PROSITE: PS00072; ACYL_COA_DH_1; 1.
 DR PROSITE: PS00073; ACYL_COA_DH_2; 1.
 DR Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;

KW

KW Mitochondrion; Transit peptide; Disease mutation; Polymorphism.
FT TRANSIT 1 33 MITOCHONDRION.
FT CHAIN 34 432 ACYL-COA DEHYDROGENASE, SHORT/BRANCHED
FT CHAIN SPECIFIC.
FT ACT SITE 414 414 BASE (BY SIMILARITY).
FT VARIANT 209 209 S -> G (in dbSNP:1799823).
FT VARIANT 255 255 /FTid=VAR_014749.
FT VARIANT 255 255 L -> F (in 2-methylbutyrylglycinuria).
FT CONFLICT 13 13 R -> K (IN REF. 3 AND 4).
FT CONFLICT 316 316 I -> V (IN REF. 3).
SQ SEQUENCE 432 AA; 47485 MW; 1EB5F894B1944E99 CRC64;

Query Match 19.28; Score 604; DB 1; Length 432;
Best Local Similarity 35.98; Pred. No. 1.4e-32;
Matches 157; Conservative 81; Mismatches 155; Indels 44; Gaps 14;

Qy 19 GLVSTANRRLLRTPPVRAFAKELFLGKIKKKEVFP-FPEVSQDE--LNEINQ--FLGP 73
Db 3 GLAV-----RLLRGRLLR-----RNFLTCLSSWKIPPHVSKSSQSEALLNTNNGIHFP 53

Qy 74 VEKFFTEVDSRK-----TDQEGKIPDELEKLSGLFLGLQVPEYGG 117
Db 54 LOTETDEEMIKSSVKKFAQBOIAPLVSTMDENSKMEKSVIQGLFQOGLMGIEVDPEYGG 113

Qy 118 LG--FSNTMYSRLGELLISMDGSITVTLAAHQAGLKGILLAGTTEQAKYLPKLASGEHI 175
Db 114 TGASPLTVLV-IEELAKVASVAFCEIQNTLTLIRKHGTEQKATYLPQLTT-EKV 171

Qy 176 AAFCLTEPASGSDAASIRSATLSEDKHYILNGSKVWITNGGLANITFTVFAKTEVVDSD 235
Db 172 GSFCLSEAGAGSDSFALKTRADKED--YYVLNGSKMWSAHEAGLFLVMANY-----D 224

Qy 236 GSVKDK-ITATIVERDFGVNGRPEDKLGIRGNTCEVHFNKIPVENILGVBGDGFK 294
Db 225 PTIGYKGTISFLVDRDTPGLHIGKPKENKLGURASSTCPLTFENVKVPANILGQIGHYK 284

Qy 295 VAMNINLGRFSMGVWAGLLKRLIEMTAETACRKFQENKRLSEFGLIQRKFAALMAQKAY 354
Db 285 YAIQSINEGRIGIAQMLGLAGQCFDYTPYIKERIQFGKRLDFDQGLQHQVAHVATQLE 344

Qy 355 VMESWYITAGMDQPGPDCSIEAMVKVPSSEAAQCSEALQILGLGTYTRDYPYER 414
Db 345 AARLLTYNAARLL-EAGKPFIK-EASMAKYAYASETAGQTTSKCIEMWGGVGTIKDYPVEK 402

Qy 415 ILRDTFRILLIPFGTNEI 431
Db 403 YFRDAKIGTIYEGASNI 419

RESULT 13
ACDS_MEGEL STANDARD; PRT; 383 AA.
AC Q06319;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2)
DE (SCAD) (Butyryl-CoA dehydrogenase) (BCAD).
OS Megaspheara elsdeni
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Acidaminococcaceae;
OC Megaspheara.
OX NCBI_Taxid=907;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 1-50.
RX MEDLINE=94002067; PubMed=8399220;
RA Becker D.F., Fuchs J.A., Banfield D.K., Funk W.D.,
RA Macgillivray R.T.A., Stankovich M.T.;
RT "Characterization of wild-type and an active-site mutant in
RT Escherichia coli of short-chain acyl-CoA dehydrogenase from
RT Megaspheara elsdeni";
RL Biochemistry 32:10736-10742(1993).
RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=95161388; PubMed=7857927;
RA Djordjevic S., Pace C.P., Stankovich M.T., Kim J.J.P.;
RT "Three-dimensional structure of butyryl-CoA dehydrogenase from
RT Megaspheara elsdeni";
RL Biochemistry 34:2163-2171(1995).
CC -!- FUNCTION: HAS AN OPTIMUM SPECIFICITY FOR 4-CARBON LENGTH FATTY
CC ACYL-COAS.
CC -!- CATALYTIC ACTIVITY: Butanoyl-CoA + ETF = 2-butenoyl-CoA + reduced
CC ETF.
CC -!- COFACTOR: FAD.
CC -!- SUBUNIT: Homotetramer.
CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; L04528; AAA03594.1; -.
DR PDB; 1BUC; 20-APR-95.
DR InterPro; IPR006088; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
DR Oxidoreductase; Flavoprotein; FAD; Fatty acid metabolism;
KW 3D-structure.
KW ACT SITE 367 367 BASE.
KW MUTAGEN 367 367 E->Q: LOSS OF ACTIVITY.
FT STRAND 2 2
FT HELIX 7 22
FT TURN 23 27
FT HELIX 28 34
FT TURN 35 35
FT HELIX 39 46
FT TURN 47 48
FT HELIX 50 52
FT TURN 53 54
FT HELIX 57 59
FT TURN 60 60
FT HELIX 62 65
FT TURN 66 66
FT HELIX 69 82
FT HELIX 84 96
FT TURN 97 97
FT HELIX 98 104
FT HELIX 107 112
FT TURN 113 113
FT HELIX 114 119
FT TURN 120 120
FT STRAND 124 127
FT STRAND 130 130
FT TURN 131 132
FT STRAND 133 133
FT HELIX 137 139
FT STRAND 143 146
FT TURN 148 149
FT STRAND 152 162
FT TURN 163 166
FT STRAND 169 175
FT TURN 182 183
FT STRAND 185 191
FT TURN 192 193
FT TURN 195 196
FT STRAND 197 202

```
FT TURN 209 210
FT STRAND 213 223
FT TURN 225 227
FT STRAND 228 229
FT TURN 232 233
FT STRAND 234 234
FT TURN 235 271
FT STRAND 273 274
FT TURN 275 276
FT STRAND 277 278
FT TURN 279 281
FT HELIX 283 312
FT TURN 313 313
FT HELIX 317 342
FT TURN 343 347
FT TURN 349 350
FT HELIX 353 360
FT TURN 361 364
FT TURN 365 367
FT HELIX 370 381
FT TURN 382 383
SQ SEQUENCE 383 AA; 3D68AAE34D9BBAB8 CRC64;

Query Match 18.8%; Score 593.5; DB 1; Length 383;
Best Local Similarity 38.1%; Pred. No. 5.8e-32;
Matches 149; Conservative 67; Mismatches 156; Indels 19; Gaps 9;

QY 54 FPPEVSQDELNEINFLGPKVEKFFTEEDSRKIDQEGKIPDTELEKLSGLGLGLOVPE 113
Db 3 FNLTDIQQDFLKLHDF---GEKKLAPTVER--DHKGIYDKELIDELLSLGTGAYFEE 57
QY 114 EYGLGFSN---TWYSRLGEIISMDGSTVTTLAAHQAIKGLIILAGTEQAKYLPKL 169
Db 58 KYGSGDDGVDVSYTLAVEELAKYDAGVAITLSATVSLCANPIWQFGTEAQKFLVPL 117
QY 170 ASGEHIAFCLTEPASGSDAASIRSRATLSEDKKHVILNGSKVWITNGLANIFVFAKT 229
Db 118 VECTKLCARGLTEPNAGTDASGOQTATKNDGT-YTLNGSKIFITNGGAADIYIVFAMT 176
QY 230 EYVSDSGSKDKTAFIVERDGGVNTGPKDKLGIRGNTCEVHFENTKIPVENILGEV 289
Db 177 D--KSGN--HGITAFILEDGTPGETYKKEKMGHTSOTMELVFQDVQKVPANLGE 232
QY 290 GDGFKVAMNLSGRFSMGVAGLLKRLIEMTAEVACTRQFNKRLSEFGLIQDKPALM 349
Db 233 GRGFKIAMTLDGGRIGVAAQALGAEALADAVEYSKORVFGPKLCKFQSISEFLADM 292
QY 350 AQKAYVMSMTYLTAGMLDQPPDCSIEAAMVKVFSSEAAWQCVSEALQILGLGYTRD 409
Db 293 KMOIEAARNLVY-KAACKKQEGK-FTVDAATAKRVASDVAMRVTTAEVQIFGGYGYSEE 350
QY 410 YPERILRDRILLIFEGTNEILRMVYALTG 440
Db 351 YPARHMRDAKITQIEGTNEVOLM---VTG 378

RESULT 14
IDV1_SOLTU STANDARD; PRT: 412 AA.
AC Q9FS88:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Isovalleryl-CoA dehydrogenase 1, mitochondrial precursor (EC 1.3.99.10)
DE (IVD 1).
GN IVD1.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=cv. Bintje; TISSUE=Tuber;
RX MEDLINE=21153238; PubMed=11231285;
RA Faivre-Nitschke S.E., Couee I., Vermel M., Grienemberger J.M.,
RA Gualberto J.M.;
RT "Purification, characterization and cloning of isovaleryl-CoA
RL dehydrogenase from higher plant mitochondria.";
RL Eur. J. Biochem. 268:1332-1339(2001).
CC -|- CATALYTIC ACTIVITY: 3-methylbutanoyl-CoA + ETF -> 3-methylbut-2-
CC enoyl-CoA + reduced ETF.
CC -|- COFACTOR: FAD.
CC -|- PATHWAY: Leucine catabolism.
CC -|- SUBUNIT: Homodimer (By similarity).
CC -|- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -|- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ278987; CAC08233.1;
DR HSP; P26440; 11VH.
DR InterPro; IPR006089; Acyl-CoA_dh.
DR InterPro; IPR006090; Acyl-CoA_dh_C.
DR InterPro; IPR006091; Acyl-CoA_dh_M.
DR InterPro; IPR006092; Acyl-CoA_dh_N.
DR Pfam; PF00441; Acyl-CoA_dh; 1.
DR Pfam; PF02770; Acyl-CoA_dh_M; 1.
DR Pfam; PF02771; Acyl-CoA_dh_N; 1.
DR PROSITE; PS00072; ACYL_COA_DH_1; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; 1.
KW Oxidoreductase; Flavo-protein; FAD; Mitochondrion; Transl. peptide.
FT TRANSIT 1 28 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 29 412 ISOVALERYL-COA DEHYDROGENASE 1.
FT ACT_SITE 273 273 BY SIMILARITY.
SQ SEQUENCE 412 AA; 45098 MW; 5DA6DF31BECE121A CRC64;
```

```
Query Match 18.6%; Score 587.5; DB 1; Length 412;
Best Local Similarity 35.5%; Pred. No. 1.6e-31;
Matches 148; Conservative 68; Mismatches 158; Indels 43; Gaps 11;

QY 38 AFARLEFLGKTKKEVFPPEVSDQELNEINQFLGPKVEKFFTEVD--SRKIDQEGKIPD 95
Db 27 AFSTSLDDTQK-----QKESVAKFAQENIAPYAEKIDRTNSFPK 68
QY 96 E-TLEKLLK-SLGLFLGLOVPEEYGGGLGFSNTMYS-RLGEIISMDGSTVTLAAHQAIKLGK 152
Db 69 EINLWKLMDGNLHGITAPEEYGGGLNGLYLYHCIALEEISRASGAVSVGVSNVCINQ 128
QY 153 IILAGTEQAKYLPKLKASGEHTAACLTEPASGSDAASIRSRATLSEDKKHVILNGSKV 212
Db 129 LVRNGTDPDQKYLKPLISGDHIGALAMSEPNAGSDVYSMKCRADRVG--GYVLNGNM 186
QY 213 WITNGLANIFTVFAPKTE-VVSDSGSVKDKTAFIVERDGGVNTGCKPEDKLGIRSNCTC 271
Db 187 WCTNGPVANTLIVYAKTDTTAGSKG-----ITAFIEKMPGFSTAKLKLKMGMSDTC 241
QY 272 EVHFENTKIPVENILGEVGGFKVAMNLSGRFSMGVAGLLKRLIEMTAEVACTRQK 331
Db 242 ELVFENCFVKNVNLGQEGGVVYVLSGLDLRLVLAAGPVGIMQACMDIVIPYVRQEQ 301
QY 332 FNKRLSEFGLIQEKF-----ALMAQKAYVMSMTYLTAGMLDQPPDCSIEAAMVKVFS 386
Db 302 FGPIGEFQLIQKLDAMTALQSSRSYVYAVAKDCDNGKIDP---KDCS----GTLLA 354
QY 387 SEAAWQCVSEALQILGLGYTRDYPYERILRDRILLIFEGTNEILRMVYALTGLQH 443
Db 355 AERATQVALQAIQCLGNGCYINEYPTGRLLRDAKAYEIAAGTSEIRRLVIGRELPRH 411
```

RESULT 15
 IVD_HUMAN
 ID IVD_HUMAN STANDARD; PRT; 423 AA.
 AC P26440; Q96AF6;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Isovalleryl-CoA dehydrogenase, mitochondrial precursor (EC 1.3.99.10)
 DE (IVD).
 GN IVD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=90203210; PubMed=2318964;
 RA Matsubara Y., Ito M., Glassberg R., Satyabhama S., Ikeda Y.,
 Tanaka K.;
 "Nucleotide sequence of messenger RNA encoding human isovaleryl-
 coenzyme A dehydrogenase and its expression in isovaleric acidemia
 fibroblasts";
 J. Clin. Invest. 85:1058-1064(1990).
 [2]
 SEQUENCE FROM N.A.
 RP MEDLINE=20143360; PubMed=10677295;
 RA Vockley J., Rogan P.K., Anderson B.D., Willard J., Seelan R.S.,
 Smith D.I., Liu W.;
 "Exon skipping in IVD RNA processing in isovaleric acidemia caused by
 point mutations in the coding region of the IVD gene";
 Am. J. Hum. Genet. 66:356-367(2000).
 [3]
 SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RC MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raja S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [4]
 SEQUENCE OF 381-423 FROM N.A.
 RP MEDLINE=93224146; PubMed=8468053;
 RA Parimoo B., Tanaka K.;
 "Structural organization of the human isovaleryl-CoA dehydrogenase
 gene";
 Genomics 15:582-590(1993).
 [5]
 ACTIVE SITE GLU-283.
 RX MEDLINE=95367542; PubMed=7640268;
 RA Mohsen A.W., Vockley J.;
 "Identification of the active site catalytic residue in human
 isovaleryl-CoA dehydrogenase";
 Biochemistry 34:10146-10152(1995).
 [6]
 X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-421.
 RX MEDLINE=97361954; PubMed=9214289;
 RA Tiffany K.A., Roberts D.L., Wang M., Paschke R., Mohsen A.W.,

Vockley J., Kim J.J.;
 "Structure of human isovaleryl-CoA dehydrogenase at 2.6-A resolution:
 structural basis for substrate specificity.";
 Biochemistry 36:8455-8464(1997).
 [7]
 VARIANTS IVA PRO-42 AND VAL-199.
 RX MEDLINE=91290048; PubMed=2063866;
 RA Vockley J., Parimoo B., Tanaka K.;
 "Molecular characterization of four different classes of mutations in
 the isovaleryl-CoA dehydrogenase gene responsible for isovaleric
 acidemia";
 Am. J. Hum. Genet. 49:147-157(1991).
 CC -!- CATALYTIC ACTIVITY: 3-methylbutanoyl-CoA + ETF = 3-methylbut-2-
 enoyl-CoA + reduced ETF.
 CC -!- COFACTOR: FAD.
 CC -!- PATHWAY: Leucine catabolism.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -!- DISEASE: Defects in IVD are the cause of isovaleric acidemia (IVA)
 [MIM:243500]. IVA is characterized by retarded psychomotor
 development, a peculiar odor resembling sweaty feet, an aversion
 to dietary protein, and pernicious vomiting, leading to acidosis
 and coma. The acute neonatal form leads to massive metabolic
 acidosis from the first days of life and rapid death.
 CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; M34192; AAA52711.1; -
 EMBL; AF191218; AAF20182.1; JOINED.
 EMBL; AF191214; AAF20182.1; JOINED.
 EMBL; AF191215; AAF20182.1; JOINED.
 EMBL; AF191216; AAF20182.1; JOINED.
 EMBL; AF191217; AAF20182.1; JOINED.
 EMBL; BC017202; AAH17202.1; -
 EMBL; AF038318; AAB92584.1; -
 EMBL; A37033; A37033.
 PIR; A37033; A37033.
 PDB; 1IVH; 20-MAY-98.
 Genew; HGNC:6186; IVD.
 GK; P26440; -
 MIM; 243500; -
 GO; GO:0008470; F:isovaleryl-CoA dehydrogenase activity; TAS.
 InterPro; IPR006089; Acyl-CoA_dh.
 InterPro; IPR006090; Acyl-CoA_dh_C.
 InterPro; IPR006091; Acyl-CoA_dh_M.
 InterPro; IPR006092; Acyl-CoA_dh_N.
 Pfam; PF00441; Acyl-CoA_dh; 1.
 Pfam; PF02770; Acyl-CoA_dh_M; 1.
 Pfam; PF02771; Acyl-CoA_dh_N; 1.
 PROSITE; PS00072; ACYL_COA_DH_1; 1.
 PROSITE; PS00073; ACYL_COA_DH_2; 1.
 Oxidoreductase; Flavoprotein; FAD; Mitochondrion; Transit peptide;
 Disease mutation; 3D-structure.
 TRANSIT 1 29 MITOCHONDRION
 CHAIN 30 423 ISOVALERYL-COA DEHYDROGENASE.
 ACT_SITE 283 283
 VARIANT 42 42
 VARIANT 199 199 L -> P (in IVA).
 MUTAGEN 283 283 /FTID=VAR_000423.
 MUTAGEN 283 283 G -> V (in IVA).
 CONFLICT 10 10 /FTID=VAR_000424.
 HELIX 38 40 E->Q; LOSS OF ACTIVITY.
 TURN 41 41 E->D; RESIDUAL ACTIVITY.
 HELIX 44 60 W -> C (IN REF. 3).

FT	TURN	61	64
ET	HELIX	65	67
FT	TURN	72	72
FT	TURN	75	76
FT	HELIX	77	87
FT	TURN	88	88
FT	TURN	90	93
FT	HELIX	96	98
FT	TURN	99	99
ET	HELIX	105	118
FT	HELIX	120	130
FT	TURN	131	133
FT	HELIX	134	140
FT	HELIX	143	155
FT	TURN	156	156
FT	STRAND	160	163
FT	STRAND	166	166
FT	TURN	167	168
FT	STRAND	169	169
FT	HELIX	173	175
FT	STRAND	179	182
FT	STRAND	186	197
FT	TURN	198	198
FT	HELIX	199	201
FT	STRAND	204	210
FT	TURN	213	214
FT	HELIX	218	221
FT	STRAND	222	228
FT	TURN	229	233
FT	TURN	232	233
FT	STRAND	234	236
FT	STRAND	242	242
FT	TURN	246	247
FT	STRAND	250	261
FT	HELIX	262	264
FT	STRAND	265	266
FT	TURN	269	270
FT	HELIX	272	287
FT	TURN	288	288
FT	HELIX	289	306
FT	TURN	307	308
FT	STRAND	310	311
FT	TURN	312	313
FT	STRAND	314	315
FT	HELIX	316	318
FT	HELIX	320	348
FT	TURN	349	350
FT	HELIX	354	379
FT	HELIX	380	384
FT	STRAND	385	385
FT	TURN	386	387
FT	TURN	389	389
FT	HELIX	390	399
FT	TURN	400	404
FT	HELIX	407	419
FT	TURN	420	420

Query Match	18.6%	Score 587.5	DB 1	Length 423
Best Local Similarity	34.1%	Pred. No. 1.7e-31		
Matches 154	Conservative 78	Mismatches 172	Indels 47	Gaps 13
QY	22	VSTANRRLL-----LRTSPVRAFAKELFLGKIKKEVPFPPEVSQDELNEINQFLGPVE	75	
Db	1	MATATRLGLGWVNASWRLRPLAGFVSQ-----RAHSLLPVDDAINGLSEEQRLRQTMA	54	
QY	76	KFFTEERV--DSRRKIDQEGKIPD--ETLEKLKSLGFLGLOVPEYVGGLGFSTWYSLRGEI	131	
Db	55	KFLQEHLPAPKAQEIDRSNEFKNLREFWKOLGNLGNLVGITAPVOYGSGGLYLEHVLVMEE	114	
QY	132	IS-MDGSITVTLLAAHQAIKGLGIIITAGTEEQRAKYLPKPLKASGSHIAAFLCTEPASGSDAA	190	
Db	115	ISRSACVAGLSYCAHNLNCINQLVRNGNEAQEKYLPKILISGYTCALAMSEPNAGSDVV	174	

Qy	191	STRSRATSEDKK---HYILNGSKVWITNGGLANIETVFAKTEVVSDSGSVKKDITAFIVE	248
Dd	175	SMLKA-----EKGKNHYYTLNGNKWFITGPDADVLIVYAKTDLAAVPAS--RCGITAFIVE	228
Qy	249	RDFGGVTNGKPKDLGRGSNTCEVHEFNTKIPVENILGEVGDSFKVANNILNSGRFSMG	308
Dd	229	KMPGFSTYSKLKDLGMRGSNCTELIEFDCKIPANILGHENGVYVLMGSLGLELRVLA	288
Qy	309	SVVAGFLRLRIEMTAEVACTRKQNKRLSBFLIQEKFA-----LMAOKAYVMESHTYLT	363
Dd	289	GGPLGLMGAOVLVDHTIPTYLHVREAFEGQIGHFQLMQCKMADMYTRLMACRQYV-----YNV	343
Qy	364	ACMLDQGPFPDCSI-EAAMVKVFSSAAWCVSSEALQILGGTGYTRYDPYERLTRIDRTL	422
Dd	344	AKACDE---GHCTAKDCAGVLYSARECATVALDGTCQFGNGNYINDFPWGRLPKDAKLY	400
Qy	423	LIFETNEILRWYIALTGLOQHAGRILTRIH	453
Dd	401	ETGATGESVRRLVI-----GRAFNADF	423

Search completed: September 3, 2003, 13:28:03
Job time : 27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2003, 13:24:57 ; Search time 42 seconds
(without alignments)
1421.923 Million cell updates/sec

Title: US-09-945-326-2
Perfect score: 3153
Sequence: 1 MSGCGLFLRTTAARACRGL.....SQQLEKRAVCAHPLDRTC 621

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3149	99.9	621	2 JC7892	acyl-CoA dehydroge
2	1356.5	43.0	655	2 A54872	acyl-CoA dehydroge
3	1308	41.5	655	2 S54183	acyl-CoA dehydroge
4	1210.5	38.4	613	2 T15905	hypothetical prote
5	821	26.0	594	2 F84085	butyryl-CoA dehydr
6	815.5	25.9	594	2 C70021	butyryl-CoA dehydr
7	711	22.5	650	2 A70817	probable fadE10 pr
8	699	22.2	382	2 B87302	acyl-CoA dehydroge
9	691.5	21.9	384	2 B83443	acyl-CoA dehydroge
10	690	21.9	379	2 G84124	acyl-CoA dehydroge
11	682	21.6	378	2 D69658	acyl-CoA dehydroge
12	682	21.6	381	2 AC3536	butyryl-CoA dehydr
13	678	21.5	381	2 B87472	acyl-CoA dehydroge
14	658	20.9	380	2 F84124	acyl-CoA dehydroge
15	652.5	20.7	379	1 S55421	acyl-CoA dehydroge
16	652	20.7	387	2 D75417	acyl-CoA dehydroge
17	648	20.6	379	2 T45286	butyryl-CoA dehydr
18	630	20.0	389	2 C70979	probable fadE25 pr
19	629.5	20.0	381	2 E75458	acyl-CoA dehydroge
20	628	19.9	456	2 AF3637	butyryl-CoA dehydr
21	626	19.9	380	2 A12983	acyl-CoA dehydroge
22	626	19.9	380	2 F98299	probable acyl-CoA
23	625.5	19.8	379	2 T47262	butyryl-CoA dehydr
24	625.5	19.8	379	2 F97233	butyryl-CoA dehydr
25	617.5	19.6	387	2 A83393	probable acyl-CoA
26	615	19.5	419	2 T15088	hypothetical prote
27	613	19.4	380	2 T44811	acyl-CoA dehydroge
28	611.5	19.4	402	2 D69530	probable acyl-CoA
29	610.5	19.4	381	2 B83791	butyryl-CoA dehydr

30 609 19.3 389 2 B87001 probable acyl-CoA
31 606 19.2 380 2 G69893 butyryl-CoA dehydr
32 604 19.2 432 1 A55680 acyl-CoA dehydroge
33 603.5 19.1 375 2 F83326 probable acyl-CoA
34 594.5 18.9 422 2 H75382 acyl-CoA dehydroge
35 593 18.8 397 2 B69378 probable acyl-CoA
36 588.5 18.7 398 2 B75411 probable acyl-CoA
37 587.5 18.6 423 2 A37033 isovaleryl-CoA deh
38 586 18.6 397 2 H84225 isovaleryl-CoA deh
39 585.5 18.6 382 2 AE3492 isovaleryl-CoA deh
40 585.5 18.6 424 1 C34252 isovaleryl-CoA deh
41 583 18.5 412 2 A30605 acyl-CoA dehydroge
42 579 18.4 412 2 B30605 acyl-CoA dehydroge
43 579 18.4 412 2 T16568 hypothetical prote
44 576.5 18.3 421 2 AD2984 acyl-CoA dehydroge
45 576.5 18.3 421 2 C98299 probable acyl-CoA

ALIGNMENTS

RESULT 1

JC7892
acyl-CoA dehydrogenase (EC 1.3.99.3) - 9 - human
C:Species: Homo sapiens (man)
C:Date: 03-Feb-2003 #sequence_revision 03-Feb-2003 #text_change 14-Apr-2003
C:Accession: JC7892
R:Zhang, J., Zhang, W.; Zou, D.; Chen, G.; Wan, T.; Zhang, M.; Cao, X.
Biochem Biophys Res Commun 297:1033-1042, 2002
A:Title: Cloning and functional characterization of hACAD9, a novel member of human a
-long-chain acyl-CoA and has specific activity toward palmitoyl-CoA (C16:0).
A:Reference number: JC7892; MUID:22246103; PMID:12359260
A:Accession: JC7892
A:Molecule type: mRNA
A:Residues: 1-621 <ZHA>
A:Cross-references: GB:AF327351
A:Experimental source: Dendritic cells
C:Comment: This enzyme, which belongs to a family of mitochondrial enzymes that catal
-long-chain acyl-CoA and has specific activity toward palmitoyl-CoA (C16:0).
C:Genetics:
A:Gene: acad-9
A:Map position: 3q26
C:Keywords: oxidoreductase

Query Match 99.9%; Score 3149; DB 2; Length 621;
Best Local Similarity 99.8%; Pred. No. 6.6e-201;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSGCGLFLRTTAARACRGLVSTANRRLLRTSPVPRAFPAKELFLGKIKKKEVFPPEVS 60
Db 1 MSGCGLFLRTTAARACRGLVSTANRRLLRTSPVPRAFPAKELFLGKIKKKEVFPPEVS 60
QY 61 QDELNEINFLGPVEKFFTEEVDSRKIDQEKIPDETLEKLSGLFLGLQVPEEYGGG 120
Db 61 QDELNEINFLGPVEKFFTEEVDSRKIDQEKIPDETLEKLSGLFLGLQVPEEYGGG 120
QY 121 SNTMYSRLGEIISMDGSIITVTLAAHQAGLKGIIILAGTEEQAKYLPKLASGEHIAAFCL 180
Db 121 SNTMYSRLGEIISMDGSIITVTLAAHQAGLKGIIILAGTEEQAKYLPKLASGEHIAAFCL 180
QY 181 TEPASGSDAASRSRATLSSEDKKHVILNGSKVWITNGGLANIFTVFAKTEVVDSDGSKVD 240
Db 181 TEPASGSDAASRSRATLSSEDKKHVILNGSKVWITNGGLANIFTVFAKTEVVDSDGSKVD 240
QY 241 KITATIVERDFGVNGKPEDKLGIRGNTCEVHENTKIPVENTLGEVGGGFKVAMN 300
Db 241 KITATIVERDFGVNGKPEDKLGIRGNTCEVHENTKIPVENTLGEVGGGFKVAMN 300
QY 301 NSGRSMGVSAGLKLRLIEMTAETACTKQFNKRLSEFGLIQEKFALMAQKAYVMSMT 360
Db 301 NSGRSMGVSAGLKLRLIEMTAETACTKQFNKRLSEFGLIQEKFALMAQKAYVMSMT 360
QY 361 YLTAGMLDQPFPPDCSIEAAMVKVFPSSAANQCVSEALQILGGGLGYTRDYPERILDR 420
Db 361 YLTAGMLDQPFPPDCSIEAAMVKVFPSSAANQCVSEALQILGGGLGYTRDYPERILDR 420

361	Db	YLTAGMDQPGPDDCSIEAAMVKVFSSEAAWOCVSEVLQILGLGTRDYPPYERILLDRTE	420
421	QY	ILLIFEGTNEILRMYYALTGLQHAGRIILTRIHELKQAKVSTVMDTVGRRLRDSLGRTV	480
421	Db	ILLIFEGTNEILRMYYALTGLQHAGRIILTRIHELKQAKVSTVMDTVGRRLRDSLGRTV	480
481	QY	LGLTGNGVHVHPSLADSANKFEENTYCFGRVTETLLLRFGKTIMEQVLKRVANILLNL	540
481	Db	LGLTGNGVHVHPSLADSANKFEENTYCFGRVTETLLLRFGKTIMEQVLKRVANILLNL	540
541	QY	YGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSQLDKYAPENLDEQIKK	600
541	Db	YGMTAVLSRASRSIRIGLRNHDHEVLANTFCVEAYLQNLFSQLDKYAPENLDEQIKK	600
601	QY	VSQOILEKRAYICAHPLDRTE	621
601	Db	VSQOILEKRAYICAHPLDRTE	621

```

RESULT 2
A54872
acyl-CoA dehydrogenase (EC 1.3.99.-) very-long-chain-specific precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Nov-1994 #sequence_revision 04-Nov-1994 #text_change 05-Nov-1999
C:Accession: A54872
R:Aoyama, T.; Ueno, I.; Kamijo, T.; Hashimoto, T.
J. Biol. Chem. 269, 19088-19094, 1994
A:Title: Rat very-long-chain acyl-CoA dehydrogenase, a novel mitochondrial acyl-CoA dehydrogenase with a
different amino acid sequence and distinct specificities of the cDNA-expressed protein.
A:Reference number: A54872; MUID:94308174; PMID:8034667
A:Accession: A54872
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-655 <AOY>
A:Cross-references: GB:D30647; NID:g5333356; PIDN:BAAO6331.1; PID:d1006895; PID:g559722
C:Keywords: oxidoreductase
E:1-40/Domain: signal sequence #status predicted <SIG>

```

Query Match	43.08;	Score	1356.5;	DB	2;	Length	655;	
Best Local Similarity	46.4%;	Pred. No.	4.7e-82;					
Matches	295;	Conservative	108;	Mismatches	192;	Indels	41;	
							Gaps	12;
Qy	9	RTTAAARACRLGVSTANRRLLR	-----TSPPVRA-----	FAKELFLGRIKKEV	53			
Db	32	RPTSAQR	-----LYASEATQAVLERPETLSSDASTREKPARAESKSAFVGMFKGQLTTDQV	87				
Qy	54	PPFVPSODELNE	-----INQFLGPVEKFFTEEDVSRKIDQEGKIPDETLEKLSLGLG	108				
Db	88	FPYPSV	-----LNEGQGTQFLKELGVFAVFEEVNDPAKNDLSKEVEEDTQGLGKELGAFG	143				
Qy	109	LQVPEEYGGFGSNMYSRLGEIISM	-DGSITVTTLAAHQAICLGKIIILAGTTEQOKAKYLP	167				
Db	144	LQVPSLGGLSNTQYARLAEIVGMHDLGVSIVLGAHQSIGFKGILLYGTKAQEKYLP	203					
Qy	168	KLASGEHTAAPCLTEPASGSDAASIRSRATLSEDKKHVILNGSKVWITNGGLANFTVFA	227					
Db	204	RVASQALAAFCLETPSSGSDVASIRSSAVSPGCKYYTLNGSKIWISSNGGLADFTVFA	263					
Qy	228	KTEVVD-SDGSVKDKITAFIVERDPGGVTNGKPKDLGIRGNTCEVHPENTKIPVENIL	286					
Db	264	KTIPIKDAATGAVKEKITAFVVERSPGGVTHGLPEKKMGIKASNTSEVYFDGVKVPDENVL	323					
Qy	287	GEVGDGFKVAMNILNSGRFSMGSVVAGLLKRLIEMTAETACTRKOFNKRSLSEFGLIOBKF	346					
Db	324	GEVGDGFKVAVNILNNGRFGMAATLAGTMKATIAKAVDHATNRTQFGQKIHNFGVIOBKL	383					
Qy	347	ALMAQAKAYVMESMTYLTAGMLDQPGFPCSTEAAMVKVFSSEAAQCSEALQIUGGLG	406					
Db	384	ARMATLQVTSMAVYMLSANMDDQ-GFKQFQIEAALISKIFGSEAAKVVTDCTQINGMGWF	442					
Qy	407	TRDYPYERILTRILLIFECTNETLRMYIALTGLQHAGRIILTRIHELKO--AKVSTVM	464					
Db	443	MKPCGAVRILDIRIFRIFEGNDILRLFVALOGCMCKGKELTGLGNALKNPLGNVGLLI	502					

```

465 DTVGRRRLRSLGRTVDLGLTGNHGVVHPISLADSNANKFEENTYCFGRVTVELLLRFSGKTIIM 524
      ::||| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::|
503 GEASKOLRRRTGTIGSGLSLS--GIHVPELSRSGELAVQALEOPATVVEAKLMKKHKGIIV 559
      :||| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::|
525 EEQLVLKRKVANIILINLYGMTAVILSRASRSIRIGLRNHDHEVLANTFCVBA---YLONLF 581
      ||| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::|
560 NEQFLLRQADGALDVLVAMVVWLSRASRSLSSEGYPPTAQEHKMLCDSWCIEAATRIRENMA 619
      :||| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::|
582 SLSQLDKYAPENLDQEIKKYSOOILEKRAYICAHPL 617
      ||| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::|
620 SLOSNPQ--OOELFRNFRSISKAWENGGLVTSNPL 653
      :||| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::| :||| |::|

RESULT 3
S54183
acyl-CoA dehydrogenase (EC 1.3.99.-) very-long-chain specific - human
C:Species: Homo sapiens (man)
C>Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 24-Sep-1998
C:Accession: S54183
R:Andresen, B.S.
submitted to the EMBL Data Library, April 1995
A:Reference number: S54183
A:Accession: S54183
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-655 <AND>
A:Cross-references: EMBL:X86556; NID:g790446; PID:g790447
C:Genetics:
A:Gene: GDB:ACADVL; VLCAD
A:Cross-references: GDB:I248185; OMIM:201475
A:Map position: 17p11.2-17p11.1
C:Keywords: oxidoreductase

```

Query Match	41.5%	Score 1308;	DB 2;	Length 655;
Best Local Similarity	46.9%	Pred. No. 7.8e-79;		
Matches	276;	Conservative 106;	Mismatches 194;	Indels 12; Gaps 8
QY	37	RAPAKELFLGKIKKKEVPEPEVSODELNE--INOFLGPVKEFFTEVDSRKLIDQEGKIPD	95	
DB	71	KSFVGMFGKGLTTDDQVPYPSVLNEEQTFKLKELPEVPSRFEFEVNDPAKNDALEMVEE	130	
QY	96	ETLEKLLGLGLGQVPEYGGGLGFSNTMYSKLGEIISM-DGSITVVTAAHQAIGLKGI	154	
DB	131	TTWQGLKELGAFGLQVPSLGGVGLCNTQYARLVEIVGMHDLGVGITLGAHOSIGFKGIL	190	
QY	155	LAGEEQAKLYPKLASEHIAAFCLTPASSDSDAASIRSRATLSEDKKHYYTLNGSKWMI	214	
DB	191	LFGTKAQKEKLYPKLASGETVAAFCLETPSSGSDAASIRSAVPSCGYIYTLNGSKLWI	250	
QY	215	TNGGLANITFTVFAKTEVDV-SDGSVKDKITAFIVERDFGGVYNGRPEDKLGIIRGNTCEV	273	
DB	251	SNGGLADITFTVPAKTPVDPATGAVKEKLTAFVVERGGGITHGPPEKKMGKIKASNTAEV	310	
QY	274	HFENKTIPIVENTILGVBGDGFKVAMNLLNGRSGMSGVVAGLLKRLIEMTAETVACTRKKQN	333	
DB	311	FDGVRVPSENVLGEVSGFKVAMHILNNGRFGMAALAGTMRGITAKAVDHATNRTQFG	370	
QY	334	KRLSEFLIQEKFALMAQKAYWESMTYLTAGMLDQPGPDCSIEAAMVYKVSSEAAWOC	393	
DB	371	EKIHNFGLLIQEKLARMVLMQLQYTESMAYMVSNMDO-GATDFQIEAAISKIFGSEAAKW	429	
QY	394	VSEALQILGGLGYTRDYPYERILRTRILLIPEGTNEILRMVYIALTGLQHAGRILTRIH	453	
DB	430	TDECIOIMGGMGFMKEPGVERVLRDLRIFRIEGTNDILRLFVALGCMCKGKELSGLS	489	
QY	454	ELKQ--AKVSTVMTDVRRLRDSLGRTVDLGTLGNHGVVHPSLADSANKFEENTYCFGR	511	
DB	490	ALKNPFEGNAGLLLGAGKQLRRRRAGLGGSLSL-GLVHPELSRSGELAVRALEQFATV	546	
QY	512	VETLLIRGCKTITWEEQVLVKRVANTILINLYGWTAVLSRASRSIRIGLRNHDHIEVLANTF	571	
DB	547	VEBAKTLKHKKGIVNEQFLLRADGDAIDYAMVVVLSRASRSLSSECHPTAOHEKMLCDTW	606	

Qy	572	CVE--AYLQNLFSLSQLDKYAPENLDEQIKKVSQIIEKRAYICAHPL	617
		: : : : : : : : : : : : : : :	
Db	607	CIEAARIREGMAALQSDPWQOE-LYRNFKSISKALYERGGVVTSNPL	653

RESULT 4

T15905
 Hypothetical protein E04F6.5 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15905
 R:Pauley, A.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of C. elegans cosmid E04F6.
 A:Reference number: Z18427
 A:Accession: T15905
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-613 <PAU>
 A:cross-references: EMBL:U28943; NID:Q861333; PID:Q861335; PIDN:AAA68357.1;
 A:experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CBSP:E04F6.5
 A:Introns: 55/3; 90/2; 195/3; 233/3; 472/2; 518/1; 567/3

Query Match	38.48;	Score 1210.5;	DB 2;	Length 613;
Best Local Similarity	43.78;	Pred. No. 2e-72;		

Matches	272;	Conservative	109;	Mismatches	205;	Indels	37;	Gaps	14;
QY	7	ELRTTAAARACRGLVSTANRRLLRSPPVRAFAKELGKIKKKEVFPPEVSQDELNE	66						
Db	16	FIRLSHAAAK-----DAKPKKAAVADSP--SFVMNFRGKAVTDQVFPYPLNMTDEQKE	69						
QY	67	-INQFLGPVEKFPTEEDSRKIDQEKIPDETLEKLSGLGLGLOVPEEYGGIGFSNTMY	125						
Db	70	TLGVMSPLEKMLVENDVVKNDSTSDIPRAVDQFAELGTGTVLPVPELEGSGFNNSOM	129						
QY	126	SRIGEII-SWDGSIYTLAAHQAIGLKGIILAGTEQKAKYKLKASGEHIAAFCLTEPA	184						
Db	130	ARVAETVGVADLFCGVVMGAHQSIGYKGLLECTDAQOKYLPDLATGRFAFAALTEPT	189						
QY	185	SGSDAASIRSRTLSEDKKHYILNGSKVITNGGLANITFVFAKTEVVDSDGSKVKITA	244						
Db	190	TGSDASSVTRAELSADGGHYVLNGGKIINSNGFADVTFVTAQTVKQADGSTDOKMSA	249						
QY	245	FIVERDFGVGTGKPKDKLIGIRGNTCEVHFENTKIPVENILGEVGDGKFAVNILNSGR	304						
Db	250	FIVERAFGCVTSPOEKKMGIKGSNTTEVHFDMLKIPVENLLGKEGEGKFAVNILNNGR	309						
QY	305	FSMGSVVAGLLKRLIEMTAIEYACTRKQFNKRISFELGIEKFALMAQKAYVMESTYTLTA	364						
Db	310	FGIPAACTGAMKHCIOKTVDHITTRVQFGKIQEFGNIOEKLVMISKLYATESIVYMLS	369						
QY	365	GMLDQPGFPDCSIEAAWVVFSEAAQCVSEALQILGLGTYTRDYPYERILRBDTRILLI	424						
Db	370	SNKDR-GIKYEIQLEAAGIKVLASEANWLVCDDAIQVHGGMFMRETGLEVRILDRIFRI	428						
QY	425	FEGTNEILRWYIALTGLQHAGRILTRIHELKQAKYSTVMDTVGRRLRDSLSGRTVDLGLT	484						
Db	429	FEGANDVLRLFLIALTGAQHAGK-----HLAEQA--SGVGGILIGLAVSRVTG-----GNT	475						
QY	485	G-NHG-VVHPSLADSANKFEENTYCFGRTVETILLRFEGKTIMEEQVLKRVANILINLYG	542						
Db	476	GSNFQGVVDASIQDSAKVLDQOIALFGQTVQGLMKHKKGIIDROYEMHRVADAANIYS	535						
QY	543	MTAVILSRASRSIRIGLRNHDHVEVLNATFCVEAYLQNLFSLSOLOKAYAPENLDEQIKKYS	602						
Db	536	SAAVLSRATYALKNKSNDAPERKVA-TYYVDKAMK---SSRFLKDGSGVENASKVAT	591						

```

QY      603 QQILEKRAYICA-----HPLD 618
        : | | : | | :
db      592 IESLAKE--VCGNGGLTLQHPVE 612

```

RESULT 5
F84085

C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: F84085
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: F84085
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-594 <STO>
 A:Cross-references: GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BA07205.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH3486

Query Match	26.0%	Score	821;	DB 2;	Length	594;			
Best Local Similarity	33.2%;	Pred. No.	1.3e-46;						
Matches	207;	Conservative	105;	Mismatches	210;	Indels	102;	Gaps	13;

QY	45	LKGIKKKEVFPPEVSQDELNEIQFGLPVEKFFT-----EEVDSRKIDQEGKIPDE	96
Db	16	LDEIDADKVF-----TPEDFTTEHMGKTTTEEFVNEVVPVLDEIENHQFD-----ISRR	66
QY	97	TLEKLKSLGLFGLQVPEEYGGGLGFSNTMYSLRGLIISMDGSIYTVLAAHQAGLKGIIIA	156
Db	67	LTUQAGEJLGLGADVPEEYGGJLGDKTISSLIITEKFSRAGGFSLSYGAHVIGISLPIVFE	126
QY	157	GTEEQAKYILPKLASGEHIAAFLCTEPASGSDAASIRSRTATLSEDKKHYYLNGSKVWITN	216
Db	127	GNEQDKKYLPLDLAGARIAAYALTPESSGSDALGAKTTAVLEAGTHYVLENGEKOWITN	186
QY	217	GGLANIFTVFAKTEVVDSDGSKOKITAFIVERDFGVTKNGKPEDKLIRGNSWTCVEHFE	276
Db	187	SAFADQFTVYAKI-----DG---BHFSAFIVEKDFEGVSTGPBEKMGKIGSRTLILE	238
QY	277	NFKIPVENILGEVGDGFKVAMNILNSRFSMGSVVACGLKRLLEMTAEVACTKPOFNKRL	336
Db	239	DALVPKENLGEVKGKHVIAFNILNVGRYKLVGVCIGGSARGIELAAKYANERKQPKPI	298
QY	337	SEFGLIQBFALMAQKAYVMESMTYLTAGMLDQ-----PGPDCGSIEA	379
Db	299	AKFTLIQEKLANMAVTVYAESSIYRTGGJFEDRLGSLSEEOQKDREYAKATAEYAI	358
QY	380	AMVKVFSSEAAWOCVSEALQTLGGLGYTRDPYPERILRDTRIILLIEGTNEILRWYIALT	439
Db	359	SLNKVVGSEALDFVADSAQVILGGYGPMAYEYRMRYRDSRINRIEFTNEINRLLPVGT	418
QY	440	GLQHA--GRI-LTTIRIHELKQAKYSTVMDTVGRRLRDSLGRTVDLGLTGNHGVVHPSLAD	496
Db	419	ILRKAMGELPFEKATALQBELMMLMPQEVG-----	450
QY	497	SANKEENTICYFGRTVETLL-----RGKTIMEEQVLKRVANILINLYGMTAVLSR	549
Db	451	-DEPLEQEKYLLSMAKVFVLMIAGTGAQTYGEKLOOEQLLANVADIVSDIFMESVILR	509
QY	550	ASRSI-RIGLRNHDHEVLLANTFCVEAY-----LQNLPFS-LSOLDK	588
Db	510	TEKAIIRNNGLEKAEQKALQVCFQEAFNRIEAAHAKESLVAMQSGDTRLTWTISILRKLTR	569
QY	599	YAPENLDEQIKKVSQQTLERKAYI	612
Db	570	HTPINVIAKKREIAGRILEAERYV	593

RESULT 6
C70021
butyryl-CoA dehydrogenase homolog vusJ - Bacillus subtilis

C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
A:Accession: C70021
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berton
C.; Bron, S.; Broutlet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Ferrar, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
y, M.; Ogawa, K.; Ogilwa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tonon, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C70021
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-594 <KUN>
A:Cross-references: GB:299120; GB:AL009126; NID:g2635613; PIDN:CAB15271.1; PID:ell84360;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yusJ

Query Match 25.9%; Score 815.5; DB 2; Length 594;
Best Local Similarity 33.4%; Pred. No. 3e-46;
Matches 201; Conservative 97; Mismatches 202; Indels 101; Gaps 12;
QY 59 VSQDELNEINFLGPEVKEFFTEEDSRKIDQEGKIPDETLEKLSGLGLGVLQVPEYGG 118
DB 45 IEQDVLPHI-----DDIENHOFHSVRL-----LKKAGELGLGADVPEYGG 88
QY 119 GFSNTMYSRLGEIISMGDSITVTAAHQATGLKGIILAGTEEQAKYLPKLASEHTAAF 178
DB 89 GLDKISSALITKEFSRAGSFSLSGAVHGIGSLPIVFGESEQKKYLPGLGASEKTAAY 148
QY 179 CLTPASGSDAASRSTRATLSEDKKHYILNCSKVITNGGLANFTVFATKEVVDSDGVS 238
DB 149 ALTEPGSSDALGAKTTAVLNEAGTHVLTGEKWIINSADVFVYAKV-----DG-- 201
QY 239 KDKITAFIVERDFGVNGKPEDKLGIRGNTCEVHFENTKIPVENILGEVGDGFKVAMN 298
DB 202 -DKTSAFIVEKEFCVGTGPEKKMGKIGSSTRLLILDQAEVPRKENVLLGEGKHVIAFN 260
QY 299 ILNSGRFSMGSVVAGLLKRLIEMTAETAYACTRQFNKRLSEFGLIOEKFALMAQKAYVNES 358
DB 261 ILNIGRYKLAVGTIGASKRVIELSAAYANQRQKFTPIAGFSLTQEKIGTWASRLYAMES 320
QY 359 MTLTAGMLDQ-----PGPPDGSIEAAMVKVFSSEAAWOCVSEALQIL 401
DB 321 SVYRTVGLFEDNMSOFTAECLKDGRQIAKSAIEVAIECSLNKVFGESETLDIVDEGVQIH 380
QY 402 GGLGYTRYDYPYERILDRPILLIPGTEIIRMYIALTGLOHA--GRI-LITRHELKQA 458
DB 381 GGYGFMQYEVYERAYDRINRIFEGTNEINRLIVPSTFLKALKGELPLPEKAQSQOE 440
QY 459 KYSTVMDTVGRRLDSGLRTDLGTGNHGVVHPSLADSANKFENTYCFGRVTETLL-- 517
DB 441 LMLMLPEEFG-----SGVL-----EQEKYIVQAKKIALFA 471
QY 518 -----RFKTIMEQLVLRVANILINLYGMTAVLSRASRSIRI-GLRNHDEHVLANT 570
DB 472 AGLAAQKYGKADREQIILVNVVINSVYAMESAVLRTKAIQAQGAQKAKQVLYTEI 531
QY 571 FCVEAYLO-----NLSLSOLDKYAPENLDEQIKKVSQQLLEKRA 610
DB 532 FVOEAFNEIEAHKESLIAMEEGDSLRMMLSLRKLTRVTPKNVIQKKREAAAGIFEAK 591
QY 611 Y 611

Db 592 Y 592

RESULT 7

A70817
probable fadE10 protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
A:Accession: A70817
R:Colet, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
C.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70817
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-650 <COL>
A:Cross-references: GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CAAL17679.1; PID:el129
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: fadE10

Query Match 22.5%; Score 711; DB 2; Length 650;
Best Local Similarity 31.4%; Pred. No. 2.9e-39;
Matches 194; Conservative 113; Mismatches 223; Indels 88; Gaps 16;
QY 38 AFAKELFLGKIKKKEVFPPEVSDDELNEINQFLGPVKEFFTEEDSRKIDQEGKIPDET 97
DB 28 SFAKELFLGRPLGLIHPFKPSDAEAEATFLVKLREFL-DTVDGVSIVIERAAQIPDEY 86
QY 98 LEKLSLGLRGLQVPEYGGIPLGFSNTMYSR-LGEIISMGDSITVTAAHQATGL-KGIIL 155
DB 87 VKGLAELGCGFLGKIPSEYGGIPLGFSNTMYSR-LGEIISMGDSITVTAAHQATGL-KGIIL 146
QY 156 AGTEEQAKYLPKLASEHTAACTLPASGSDAASRSTRATLSEDKKHYILNCSKVIT 215
DB 147 AGTAEOKKRRFLPCAAGA-ISAFLLTEPDGSDPARNASTATPIDDQQAQVEGKLVTT 205
QY 216 NGGLANFTVFATKEVVDSDGSKDKITAFIVERDFGVNGKPEDKLGIRGNTCEVHF 275
DB 206 NGVVADLLVMAR--VPRSEGH-RGGISAFVVEADSPGITVERNRKPMGLRGIENGVTRL 262
QY 276 ENTKIPVENILGEVGDGFKVAMNILNSGRFSMGSVVAGLLKRLIEMTAETAYACTRQFNK 335
DB 263 HRVVPKDNLTIGREGDGLKIALTLNAGRLSLPAIATGVAKQALUKIAREWSVERVQMGKP 322
QY 336 LSEFGLIOEKFALMAQKAYVNESMTYLTAGMDQPGPPDGSIEAAMVKVFSSEAAWOCVS 395
DB 323 VQGHAEVASKISFIATNYALDAVVELSSQMADE-GRNDFIEAALAKLNSSEMACLVDG 381
QY 396 EALQILGGLGY-----TRDYPYERILDRPILLIPGTEIIRMYIALTG-----L 441
DB 382 ELIQRIGRGVYETAESLAARGERAVPVQMVRLRINRIFEGSSEIMRLIAREAVDAHL 441
QY 442 QHAGRILTRHELKQAKVSTVMTVGRRLDSGLRTDLGTGNHGVHPSLA-----495
DB 442 TAAGDLNPKA-DLQKAAAAA-----GASGYAKWLPKLVFEGEQL 482
QY 496 -----DSANKFENTYCFGRVTETLLLRFRGKTIMEQLVLRVANIL 537
DB 483 PTTYREFCALATHLRFVRSRKLARTIF-YG-----MARWQASLEKKQKQGLGRIVDIG 535
QY 538 INLYGMTAVLSRASRSIRI-GLRNHDEHVLANTFC-----VEAYLONLFSLSOLDKYAP 591
DB 536 AELFAISACVRAEQRTADPVEGEQAYELAEAFCCQOATLREALFDALWS-----586
QY 592 ENLDEQIKKVSQQLLEK 609
DB 587 -NTSDIVRLANDVLEGR 603

RESULT 8

B87302
acyl-CoA dehydrogenase, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: B87302
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4138-4141, 2001
A:Title: Complete genome sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87302
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <STO>
A:CROSS-references: GB:AE005673; NID:g13421592; PIDN:AAK22414.1; GSPDB:GN00148
C:Genetics:
C:Superfamily: acyl-CoA dehydrogenase

Query Match 22.2%; Score 599; DB 2; Length 382;
Best Local Similarity 38.5%; Pred. No. 8.2e-39;
Matches 155; Conservative 80; Mismatches 124; Indels 44; Gaps 8;

Qy 63 ELNEINQFLGPKVEKFFTEEDV--SRKIDQEGKIPDETLEKLSGLFGLQVPEEYGGGLG- 120
Db 4 DLETREQIDTVARFVAERLRPIEAQVAENDVDDVIEEMKGLGLGLTIPEEFGGLG- 62
Qy 121 SNTWYSLRGEIISMDGSIYTLAAHQ-----IGLKGIILAGTEEGKAKYLP 167
Db 63 -----LTMEERALVAIELGRASPAPRSVFGTVNGVIGSGGLVMEFNDQKAKWLP 111
Qy 168 KLASGEHIAFCLTEPAGSGDAASIRSRATLSEDKKHYILNGSKVWTNGLANIFTVFA 227
Db 112 GIAGGAVITFALTEPAGSGDAASVQTRAT--RDGDDYILNGSKRYITNAGKASLETVMA 169
Qy 228 KTEVVDSDGSKVKDKITAFIVERDFGVGNGKPKDLGIRGNTCEVHFENTKIPVENILG 287
Db 170 RTN-PDAKGA--GVSAFLVPRDLPLGTVGKPKKQGAHLDHVTDFNVRVPMNRLG 226
Qy 288 EVGDGFKVANNILNGSRFSGVAVGLLKRLEIEMTAETACTRKQFNKRLSEFGLIQ---- 343
Db 227 AEGEGFKVAVQVLDGRGLHIAACVGVAAERLIADCVAYASERKQFGOPIASFOLIQAIA 286
Qy 344 -EKFALMAOKAYVMESMTYLTAGMLDQPGFPCDSIEAAMVKVFSSEAAMQCVSEALQILG 402
Db 287 DSTELAAKALVLETARKRDAGV-----NVTLEAASKLPFASEMVGVRADRAVQVFG 339
Qy 403 GLGYTRDYPYERILDRTRILLIFEGTNEILRMVIALTGLQHAG 445
Db 340 GAGVADYGIERLYRDVRIEYEGTSQVQLLIARETLKRG 382

RESULT 9

B83443
probable acyl-CoA dehydrogenase PA1631 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83443
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83443
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <STO>
A:CROSS-references: GB:AE004590; GB:AE004091; NID:g9947587; PIDN:AAG05020.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA1631

C:Superfamily: acyl-CoA dehydrogenase

Query Match 21.9%; Score 691.5; DB 2; Length 384;
Best Local Similarity 37.8%; Pred. No. 2.6e-38;
Matches 147; Conservative 82; Mismatches 121; Indels 39; Gaps 8;

Qy 67 INOFLGPKVEKFFTEEV--DSRKIDQEGKIPDETLEKLSGLFGLQVPEEYGGGLGFSNTM 124
Db 8 LNLLEDSDIRQFVRESLVPHEQVAETDRIPEAIIRARMREGLFGLSIPAYGGGLG----- 62
Qy 125 YSLRGEIISMDGSIYTVT-----LAAHQALGLKGIILAGTEEGKAKYLPKLAS 171
Db 63 -----VTMEEEVSIAPFELGRTPAPRSLTLGTTNNGIGSQGIVIDGTEQKRLYPLRAS 115
Qy 172 GEHIAAFCLTEPAGSGDAASIRSRATLSEDKKHYILNGSKVWTNGLANIFTVFAKTEV 231
Db 116 GELLSSFLTEPDSGSDAASLKTAV--RGEHYVLNGTKRFTTNAPOAGIYVWART-- 171
Qy 232 VDSGSKVKD--ITAFIVERDFGVGTVNGKPKDLGIRGNTCEVHFENTKIPVENILGEY 289
Db 172 ---DPAIRGAGGISAFVVERGTPGLSLGKPKDRKMGOKGAHTCDVIFDCRPASQLIGV 228
Qy 290 -GDGEFKVANNILNGSRFSGVAVGLLKRLEIEMTAETACTRKQFNKRLSEFGLIOEKFAL 348
Db 229 EGVEFKTAMKVLQDGRHLIAACVGVAAERLMDALRYALERKQFGOPIAEFQLIQAMLAD 288
Qy 349 MAQKAYVMESMTYLTAGMLDQPGFPCDSIEAAMVKVFSSEAAMQCVSEALQILGGLGYTR 408
Db 289 SKAEYAAARCMVIDAARQRDEG--RDVGTASCAKLPFASEMCGRVADRAVQIFGGAGYIG 346
Qy 409 DYPYERILDRTRILLIFEGTNEILRMVIA 437
Db 347 DYGIERYDVRLEFRIVEGTTIQQLLIA 375

RESULT 10

G84124
acyl-CoA dehydrogenase mmgC [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G84124
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G84124
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-379 <STO>
A:CROSS-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07518.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
C:Superfamily: acyl-CoA dehydrogenase

Query Match 21.9%; Score 690; DB 2; Length 379;
Best Local Similarity 40.8%; Pred. No. 3.2e-38;
Matches 147; Conservative 74; Mismatches 125; Indels 14; Gaps 6;

Qy 78 FTEEVDSRKIDQEGKIPDETLEKLSGLFGLQVPEEYGGGLGFSNTMY-SRLGEIISMDG 136
Db 27 FVEEMET-----NDTFPMHIVKEMGELGUMGIPIPEAYGGAEMDMFTSYIIAHELSKVA 81
Qy 137 SITVTLLAAHQALGLKGIILAGTEEGKAKYLPKLASGEHIAAFCLTEPAGSGDAASIRSR 196
Db 82 TVGVILSVHTSVGTNPILFEGTEEQSKRIPKLAKGEYLCAFLTEPAGSGDAANLKTTA 141
Qy 197 TLSEDKKHYILNGSKVWTNGLANIFTVFAKTEVVDSDGSKVKDKITAFIVERDFGGVTN 256
Db 142 LKQGD--HYRLNGSKVFTITNGGAADTVVFAST----NPSAGRKGISAFIVEKETPGFTV 195

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: B87472

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-381 <STO>

A:Cross-references: GB:AE005673; NID:g13423230; PIDN:AAK23774.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC1798

C:Superfamily: acyl-CoA dehydrogenase

Query Match 21.5%; Score 678; DB 2; Length 381;
Best Local Similarity 38.3%; Pred. No. 2e-37;

Matches 149; Conservative 81; Mismatches 135; Indels 24; Gaps 7;

QY 67 INQFLGPVEKFFTEVDNRK--IDQEKIPDETLEKLSGLGLQVPEEYGGGLGFSNTM 124

Db 7 LSALIDVIOFVAERLRPIEGVSETDEVPSIIEEMKQLGLFSLPESVGGGLGSLLEE 66

QY 125 YSRLEGIISM--DGSITVTLLAHOAIGLKGIIAGTTEQKAKYLPLKLSAGEHIAFCLT 181

Db 67 EARV--IVAFCHTAPAFRSTFTGNVIGISQGLVMFGDEAQAARWLPSIASGETITAFALT 124

QY 182 EPASGSDAASIRSRATLSDEKHHVILNGSKVWITNGGLANFTVFAKTEVVDSDGSKDK 241

Db 125 EAEAGSDASVQTRAV--RQDGHVNLGVKRYIITNAGRANLFTVMARTDPTNKGA---G 179

QY 242 ITATIVERDFGGVTNGKPEDKLGIRGNTCEVHPENTKIPVENILGEVGDGFKVAMNIN 301

Db 180 VSAFLVPADLPGLSVGPKMKMGOGGAHIDVVFEDVRVPENRLGAEAGGFTVAMRVLD 239

QY 302 SGRFSMGSVAGLLKRLIEMTAETACTRKQFNKRLSEFLIQ-----EKFALMAQAAYVM 356

Db 240 RGRVHISAVCVGVAERLIADCVAASERKQFGOPISFQITQAMIAADSKTEALAAKALVF 299

QY 357 ESMYTLRAGMDQFGFPDPCSTEAAMVKVFSSEAAQVCSEALQTLGLGYTRDYPYERIL 416

Db 300 DTARKRDAG-----ANVTLEAAATKLFASEMGRVADRAVQVFGAGVADYGIERYL 352

QY 417 RDTIRILLIFGTNEILRMVYIALTGLQHAQ 445

Db 353 RQVRIFRYEGASQIQIQLIARETLKRG 381

RESULT 14

F84124

acyl-CoA dehydrogenase acdA [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: F84124

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: F84124

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <STO>

A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07517.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: acdA

C:Superfamily: acyl-CoA dehydrogenase

Query Match 20.9%; Score 658; DB 2; Length 380;

Best Local Similarity 40.7%; Pred. No. 4.2e-36;

Matches 156; Conservative 66; Mismatches 139; Indels 22; Gaps 8;

QY 62 DELNEINQFLGPVEKFFTEEV--DSRKKIDQEKIPDETLEKLSGLGLQVPEEYGGGLG 119

Db 7 DEQQMIKK--VROFAENAVPTAEERDEERDFRFDQMAEGLGIPWPEYGGIG 63

QY 120 FSNWY-SRLGHIISMOGSITVTLLAHOAIGLKGIIAGTTEQKAKYLPLKLSAGEHIAAF 178

Db 64 ADYVSYVIAVEELSRVCASTGVTLTSAHLSASWPMYKFGTEQKQOFLRPMABCKKIGAY 123

QY 179 CLTEPASGSDAASIRSRATLSDEKHHVILNGSKVWITNGGLANFTVFAKTEVVDSDGVS 238

Db 124 GLTEPGSGSDAANKTKTAVLEGD--DYILNGSKLIFTNGGIADIYIVFAVTE----- 173

QY 239 KDK-----ITAFIVERDFGGVTNGKPEDKLGIRGNTCEVHPENTKIPVENILGEVGDGFK 294

Db 174 PEKRHKGVSAFIVEADTPGFSVGKKKKLGRSPPTTEILFEIDCRVPKKNLCKEGDGPK 233

QY 295 VANNILNMGFSMGSVAGLLKRLIEMTAETACTRKQFNKRLSEFLIQEKFALMAQAAY 354

Db 234 IANMTLDGGRNGTAAGAVGIAQGLDAVAAYANERKQFGKPGICQGGIAPFKLADMATKVE 293

QY 355 VMESMYTLRAGMDQFGFPDPCSTEAAMVKVFSSEAAQVCSEALQTLGLGYTRDYPYER 414

Db 294 ASRLTYT-QAAMRESEGL-SYKESAMSKLFGAGDTAMDVTVEAVQVFGGYGTYKTPVER 351

QY 415 ILRDTIRILLIFGTNEILRMVYIA 437

Db 352 YMRDAKITQIYEGTNEIQLRVIS 374

RESULT 15

S55421

acyl-CoA dehydrogenase (EC 1.3.99.3) acdA - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C:Accession: S55421; B69581

R:Glaser, P.; Danchin, A.

A:Description: Cloning and sequencing of the Bacillus subtilis chromosomal region fro

submitted to the EMBL Data Library, May 1995

A:Reference number: S55414

A:Accession: S55421

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-379 <GLA>

A:Cross-references: EMBL:249782; NID:g853752; PIDN:CAA89868.1; PID:g853760

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabel, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y.M.; Ogawa, K.; Ogiwara, C.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togononi, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: B69581

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-379 <KUN>

A:Cross-references: GB:299123; GB:AL009126; NID:g2636240; PIDN:CAB15745.1; PID:g26362

A:Experimental source: strain 168

C:Genetics:

A:Gene: acdA

C:Superfamily: acyl-CoA dehydrogenase

C:Keywords: oxidoreductase

Query Match 20.7%; Score 652.5; DB 1; Length 379;
Best Local Similarity 42.1%; Pred. No. 9.8e-36;

Matches 155; Conservative 60; Mismatches 140; Indels 13; Gaps 7;

QY 74 VEKFFTEEV--DSRKKIDQEKIPDETLEKLSGLGLQVPEEYGGGLGFSNTWY-SRLGE 130

Db 16 VRDAKHEVAPTAAERDEQERDFREMANLGLTGPWPEDYGGIGSYLAYVIAVEE 75

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 10:14:14 : Search time 6852.73 Seconds
(without alignments)
11121.789 Million cell updates/sec

Title: US-09-945-326-3
Perfect score: 1863
Sequence: 1 atgagcgtcggtggtctt.....accctctggacaggacatgc 1863

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
- 1: gb_ba:*
 - 2: gb_htg:*
 - 3: gb_in:*
 - 4: gb_om:*
 - 5: gb_ov:*
 - 6: gb_pat:*
 - 7: gb_ph:*
 - 8: gb_pl:*
 - 9: gb_pr:*
 - 10: gb_ro:*
 - 11: gb_sts:*
 - 12: gb_sy:*
 - 13: gb_un:*
 - 14: gb_vl:*
 - 15: em_ba:*
 - 16: em_fun:*
 - 17: em_hum:*
 - 18: em_in:*
 - 19: em_mu:*
 - 20: em_on:*
 - 21: em_or:*
 - 22: em_ov:*
 - 23: em_pat:*
 - 24: em_ph:*
 - 25: em_pl:*
 - 26: em_ro:*
 - 27: em_sts:*
 - 28: em_un:*
 - 29: em_vl:*
 - 30: em_htg_hum:*
 - 31: em_htg_inv:*
 - 32: em_htg_other:*
 - 33: em_htg_mus:*
 - 34: em_htg_pln:*
 - 35: em_htg_rod:*
 - 36: em_htg_mam:*
 - 37: em_htg_vrt:*
 - 38: em_sy:*
 - 39: em_htgo_hum:*
 - 40: em_htgo_mus:*
 - 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1863	100.0	1863	6	AX394749	AX394749 Sequence
2	1863	100.0	2448	9	BC013354	BC013354 Homo sapi
3	1863	100.0	2452	6	AX394747	AX394747 Sequence
4	1859.8	99.8	2399	6	BD158773	BD158773 Primer fo
5	1859.8	99.8	2399	9	AK024012	AK024012 Homo sapi
6	1859.8	99.8	2494	9	BC007970	BC007970 Homo sapi
7	1858.2	99.7	2411	9	AF327351	AF327351 Homo sapi
8	1671.4	89.7	2400	6	BD157796	BD157796 Primer fo
9	1671.4	89.7	2400	9	AK022568	AK022568 Homo sapi
10	1380.8	74.1	2129	10	BC033277	BC033277 Mus muscu
11	1380.8	74.1	2148	10	BC032213	BC032213 Mus muscu
12	1380.8	74.1	3071	10	BC031137	BC031137 Mus muscu
13	1029	55.2	1494	9	BC001817	BC001817 Homo sapi
14	704.6	37.8	771	6	BD150623	BD150623 primer fo
15	526	28.2	1995	9	AK027186	AK027186 Homo sapi
16	499.6	26.8	772	6	BD129328	BD129328 Human gen
17	496.4	26.6	544	6	BD109621	BD109621 EST and e
18	484.4	26.0	543	6	AX588711	AX588711 Sequence
19	460.4	24.7	507	6	AX588841	AX588841 Sequence
20	418.4	22.5	431	6	AX072562	AX072562 Sequence
21	396.4	21.3	2058	9	AB083302	AB083302 Macaca fa
22	396.4	21.3	2103	10	MTMVLCAD	271189 M.musculus
23	396	21.3	2140	4	BTU30817	U30817 Bos taurus
24	394.8	21.2	2171	10	BC026559	BC026559 Mus muscu
25	393.2	21.1	796	6	BD147321	BD147321 Primer fo
26	391.6	21.0	1965	6	BD141736	BD141736 IRAP bind
27	391.6	21.0	2168	9	BC020218	BC020218 Homo sapi
28	391.6	21.0	2169	9	HSHVLCAD	D43682 Human mRNA
29	391.6	21.0	2219	9	HSHVLCAD	X86556 H.sapiens H
30	391.6	21.0	2231	9	BC012912	BC012912 Homo sapi
31	391.6	21.0	2258	9	AK056083	AK056083 Homo sapi
32	391.6	21.0	2259	9	BC000399	BC000399 Homo sapi
33	391.2	21.0	1840	10	AF017176	AF017176 Mus muscu
34	388.6	20.9	2117	6	AX401908	AX401908 Sequence
35	388.6	20.9	2117	10	RATVLCAD	D30647 Rat mRNA fo
36	383	20.6	1439	10	AY292988	AY292988 Rattus no
37	380.2	20.4	434	6	AX071059	AX071059 Sequence
38	372.8	20.0	171981	3	AC008344	AC008344 Drosophil
39	372.8	20.0	176813	3	AC008096	AC008096 Drosophil
40	372.8	20.0	176980	2	AC020290	AC020290 Drosophil
41	372.8	20.0	277421	3	AE003796	AE003796 Drosophil
42	372.6	20.0	412	6	AX071159	AX071159 Sequence
43	339.8	18.2	2027	3	AK115332	AK115332 Ciona int
44	283.2	15.2	300	6	BD129040	BD129040 Human gen
45	283	15.2	1476	9	AK058109	AK058109 Homo sapi

ALIGNMENTS

RESULT 1
AX394749
LOCUS AX394749
DEFINITION Sequence 3 from Patent WO0218582.
ACCESSION AX394749
VERSION AX394749.1 GI:21065828
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Meyers, R. and Hunter, J.J.
TITLE 62112, a novel human dehydrogenase and uses thereof
JOURNAL Patent: WO 0218582-A 3 07-MAR-2002;

Millennium Pharmaceuticals, Inc. (US)		
FEATURES	Location/Qualifiers	
source	1. .1863	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
CDS	1. .>1863	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD32050.1"	
	/db_xref="GI:21065829"	
	/translation="MSGCGFLRTTAAARACRLVSTANRLLRTSPVRAFAPKELF	
	LGRIKKVEFPFVQSDELNEINQFLPVEFFTEEDSRKIDQEGKIPDETLEKLK	
	SLGLFQLOPVEEYGGFSNTMYSLRGEIISMDGSITVTAAHQAILGKIIILAGTEE	
	OKAKYLPKLAGSHIAAFCITPEASGSDAASIRSRATLSDEKKHYILNGSKVITNGG	
	LANIFVFAKTEVSDGSVKDKITAFIVERDPGGVNGKPKDKLIGIRSNCTCEVHFE	
	NTKIPVENILGVDGDKVAMNLLNSGRFSWGSVYAGLLKRLLEWTAEVACTRKOPNK	
	RUSEFGLIOEKFALMAOKAYVMESMTYLTAGMDIDPQFPDCEAEAMKVFSEAAWQ	
	CVSEALQILGLGYTRDYPYERILRDRITRILLIPEGTNEILRMVYIALTGOHAGRILTT	
	RIHELQAKVSTVMDTVGRRLRSLGRITDLGLTGNHVVHPSLADSANKFEENTYCF	
	GRVETLLLRFGVTIMEEOLVLRVANILINLYGMTAVLSRASRSIRIGLRNHDHVL	
	LANTFCVEAYLQNLFSLSQLDKVPENLDEQIKKVSQIILEKRAYICAHPLDRTC"	
BASE COUNT	471 a 481 c 519 g 392 t	
ORIGIN		
Query Match	100.0%; Score 1863; DB 6; Length 1863;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1863; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATGAGCGGCTCGGGCTCTTCCTGCGGACACACGCGTCGCGCTCGTGCCTGCCGGGCTG 60	
Db	1 ATGAGCGGCTCGGGCTCTTCCTGCGGACACACGCGTCGCGCTCGTGCCTGCCGGGCTG 60	
Qy	61 GTGTCTCTACCGGAAACCGCGGCTACTCGGACACACGCGGCTGTACGAGCTTTGCGC 120	
Db	61 GTGTCTCTACCGGAAACCGCGGCTACTCGGACACACGCGGCTGTACGAGCTTTGCGC 120	
Qy	121 AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAAGTTTTCCCATTTCCAGAAAGTTAGC 180	
Db	121 AAAGAGCTTTTCTTAGGCAAAATCAAGAAGAAAGTTTTCCCATTTCCAGAAAGTTAGC 180	
Qy	181 CAAGATGACCTTAATGAATCAATCAGTCTCTGGGACCCGCTGGGAAATTTCTCAGTGAA 240	
Db	181 CAAGATGACCTTAATGAATCAATCAGTCTCTGGGACCCGCTGGGAAATTTCTCAGTGAA 240	
Qy	241 GAGTGGACTCCCGGAAATTTGACAGGAGGAAAAATCCAGATGAATTTGGAGAAA 300	
Db	241 GAGTGGACTCCCGGAAATTTGACAGGAGGAAAAATCCAGATGAATTTGGAGAAA 300	
Qy	301 TTGAAGACCTTAGGGCTTTTGGGCTCAAGTCCAGAGAAATATGCTGGGCTGCGGCTTC 360	
Db	301 TTGAAGACCTTAGGGCTTTTGGGCTCAAGTCCAGAGAAATATGCTGGGCTGCGGCTTC 360	
Qy	361 TCCAACCATGTACTCAAGACTAGGAGAGATCATCAGCATGGATGGGTCCATCAGTGTG 420	
Db	361 TCCAACCATGTACTCAAGACTAGGAGAGATCATCAGCATGGATGGGTCCATCAGTGTG 420	
Qy	421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGAAGAG 480	
Db	421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGAAGAG 480	
Qy	481 CAGAAAGCAATATCTTCCCTAAATGCGTCCGGGAGACATTTGACGCTTCTGCTC 540	
Db	481 CAGAAAGCAATATCTTCCCTAAATGCGTCCGGGAGACATTTGACGCTTCTGCTC 540	
Qy	541 ACGGAGCCAGCCAGTGGAGCGATGCAGCCTCAATTCGGAGCAGAGCCACACTAAGTGAA 600	
Db	541 ACGGAGCCAGCCAGTGGAGCGATGCAGCCTCAATTCGGAGCAGAGCCACACTAAGTGAA 600	
Qy	601 GACAGAGAGCACTACATCCTCAATGGCTCCAGGTCTGGATTACTAATGAGGAGTGGCC 660	
Db	601 GACAGAGAGCACTACATCCTCAATGGCTCCAGGTCTGGATTACTAATGAGGAGTGGCC 660	
Qy	661 AATATTTTACTGTGTTTGCRAAGACTGAGGTCTGTTGATTCTGTGATGGATGGAAGAC 720	
Db	661 AATATTTTACTGTGTTTGCRAAGACTGAGGTCTGTTGATTCTGTGATGGATGGAAGAC 720	
Qy	721 AAAATCACAGCATTATAGTAGAAAAGAGACTTTTGGTGAGTCACTAAATGGGAAACCCGAA 780	
Db	721 AAAATCACAGCATTATAGTAGAAAAGAGACTTTTGGTGAGTCACTAAATGGGAAACCCGAA 780	
Qy	781 GATAAATTAGGCATTTCGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGATA 840	
Db	781 GATAAATTAGGCATTTCGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGATA 840	
Qy	841 COTGTGAAAAACATCCTTGCAGAGTTCGAGATGGTAAAGTGGCCATGAACATCCTC 900	
Db	841 COTGTGAAAAACATCCTTGCAGAGTTCGAGATGGTAAAGTGGCCATGAACATCCTC 900	
Qy	901 AACAGCGGCGGTTCAGCATGGGACGCTGCTGGTGGGCTGCTCAAGAGATTGATGAA 960	
Db	901 AACAGCGGCGGTTCAGCATGGGACGCTGCTGGTGGGCTGCTCAAGAGATTGATGAA 960	
Qy	961 ATGACTGCTGAGTACGCTGCACAGGAAACAGTTTAAACAAGAGGCTCAGTGAATTTGA 1020	
Db	961 ATGACTGCTGAGTACGCTGCACAGGAAACAGTTTAAACAAGAGGCTCAGTGAATTTGA 1020	
Qy	1021 TTGATTTCAGGAAATTTTGCACCTGATGGCTCAGAAGCTTACGTCTATGGAGATGATGACC 1080	
Db	1021 TTGATTTCAGGAAATTTTGCACCTGATGGCTCAGAAGCTTACGTCTATGGAGATGATGACC 1080	
Qy	1081 TACCTCACAGCAGGATGCTGGACCAACCTGGCTTTCCGAGCTGCTCATCAGGAGCGC 1140	
Db	1081 TACCTCACAGCAGGATGCTGGACCAACCTGGCTTTCCGAGCTGCTCATCAGGAGCGC 1140	
Qy	1141 ATGGTGAAGGTGTTTACGCTCCGAGCGGCTGCGAGTGTGAGTGAGCGCTGACAGATC 1200	
Db	1141 ATGGTGAAGGTGTTTACGCTCCGAGCGGCTGCGAGTGTGAGTGAGCGCTGACAGATC 1200	
Qy	1201 CTCGGGGCTTGGGCTACACAAGGAGCTATCCGTACGAGCGCATCTGCGTGACACCCGC 1260	
Db	1201 CTCGGGGCTTGGGCTACACAAGGAGCTATCCGTACGAGCGCATCTGCGTGACACCCGC 1260	
Qy	1261 ATCCTCCTCATCTTCGAGGAAACCAATGAGATTCTCGGATGTACATGCCCTGAGCGGT 1320	
Db	1261 ATCCTCCTCATCTTCGAGGAAACCAATGAGATTCTCGGATGTACATGCCCTGAGCGGT 1320	
Qy	1321 CTCAGCATGCGCGCGCATCTCTGACTTACCAGGATCCATGAGCTTAAACAGGCGCAAGTG 1380	
Db	1321 CTCAGCATGCGCGCGCATCTCTGACTTACCAGGATCCATGAGCTTAAACAGGCGCAAGTG 1380	
Qy	1381 AGCAGAGTCAATGATACCGTTGGCGGAGCTTTCGGGACTTCCCTGGGCGCACTGTGGAC 1440	
Db	1381 AGCAGAGTCAATGATACCGTTGGCGGAGCTTTCGGGACTTCCCTGGGCGCACTGTGGAC 1440	
Qy	1441 CTGGGCTGACAGCAACCAATGAGTTGTGACCCAGCTTCTCGGACAGTGCACAACAG 1500	
Db	1441 CTGGGCTGACAGCAACCAATGAGTTGTGACCCAGCTTCTCGGACAGTGCACAACAG 1500	
Qy	1501 TTTGAGGAGAACACCTACTGCTTTCGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1560	
Db	1501 TTTGAGGAGAACACCTACTGCTTTCGGCGGACCGTGGAGACACTGCTGCTCCGCTTTGGC 1560	
Qy	1561 AAGACCATCATGAGGAGCAGCTGTTACTAAGCGGTGGCCAACTCCTCATCAACCTG 1620	
Db	1561 AAGACCATCATGAGGAGCAGCTGTTACTAAGCGGTGGCCAACTCCTCATCAACCTG 1620	
Qy	1621 TATGGCATGACGGCGCTGCTGTCGGGCGCAGCGCTTCCATCCGCAATTTGGGCTCCGCAAC 1680	
Db	1621 TATGGCATGACGGCGCTGCTGTCGGGCGCAGCGCTTCCATCCGCAATTTGGGCTCCGCAAC 1680	
Qy	1681 CACGACCAAGGCTCTCTTTGGGCAACACCTTCTGCGTGGAAAGCTTACTTGCAGAAATCTC 1740	
Db	1681 CACGACCAAGGCTCTCTTTGGGCAACACCTTCTGCGTGGAAAGCTTACTTGCAGAAATCTC 1740	
Qy	1741 TTGACGCTCTCTCAGCTGGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAA 1800	
Db	1741 TTGACGCTCTCTCAGCTGGACAGTATGCTCCAGAAACCTAGATGAGCAGATTAAGAAA 1800	

```

|||||
1741 TTACGCCCTCTCTCAGCTGGACAAGTAGTCCAGAAAACCTAGATGACGAGATTAGAAA 1800
QY 1801 GTGTCCCGCAGAGATCCTTTGAGACCGAGCCTATATCTGTGCCACCCCTCGGACAGGACA 1860
Db 1801 GTGTCCCGCAGAGATCCTTTGAGACCGAGCCTATATCTGTGCCACCCCTCGGACAGGACA 1860
QY 1861 TGC 1863
Db 1861 TGC 1863

RESULT 2
BC013354
LOCUS
DEFINITION
  Homo sapiens very-long-chain acyl-CoA dehydrogenase VLCAD, mRNA
ACCESSION
  BC013354
VERSION
  1
KEYWORDS
  MGC.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 2448)
  Strausberg,R.L., Feingold,E.A., Grouse,L.H., DeRge,J.G.,
  Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
  Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
  Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Bhat,N.K.,
  Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
  Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
  Scheetz,T.E., Brownstein,M.J., Ussdin,T.B., Toshiyuki,S.,
  Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
  Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
  McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
  Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
  Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
  Sanchez,A., Whiting,M., Madan,A., Madan,A., Rodriguez,S.,
  Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
  Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
  Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,
  Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  Generation and initial analysis of more than 15,000 full-length
  human and mouse cDNA sequences
  Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  22388257
  12477932
  2 (bases 1 to 2448)
  Strausberg,R.
  Direct Submission
  Submitted (31-AUG-2001) National Institutes of Health, Mammalian
  Gene Collection (MGC), Cancer Genomics Office, National Cancer
  Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
  USA
  NIH-MGC Project URL: http://mgc.nci.nih.gov
  Contact: MGC help desk
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: DCTD/DTF
  CDNA Library Preparation: Rubin Laboratory
  CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Institute for Systems Biology
  http://www.systemsbio.org
  contact: amadan@systemsbiology.org
  Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha
  Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
  Clone distribution: MGC clone distribution information can be found
  through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
  Series: IRAL Plate: 25 Row: c Column: 21.
  Location/Qualifiers
    1..2448
    /organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_mgc="14970 IMAGE:3935925"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH_MGC_7"
/lab_host="DH10B-R"
/Note="Vector: pOTB7"
1..2448
/gene="VLCAD"
/Note="synonyms: NP0002, MGC14452, ACAD-9"
/db_xref="LocusID:28976"
66..1931
/codon_start=1
/product="very-long-chain acyl-CoA dehydrogenase VLCAD"
/protein_id="AAH13354.1"
/db_xref="GI:15426498"
/db_xref="LocusID:28976"
/translation="MSGGGLFRTTAARACRGLVSTANRRLLRTSPVRAFAKELF
LGKIKKKEVFPEVSDQLNEINQFGLPVKFFTEVDNRKIDQEKIPDETLEKLG
SLGFLQVPEEYGGGLFNTMYSLRGLIISMDGSIITVLAHQAIKGLIILAGTGE
QKAKYLPKLSAGEHIAAPCLTEPASGSDAASIRSRATLSDDKHVILNGSKVITNGG
LANITFTFAKTEVVDSDGVKDKITAFIVERDFGVGTNGKPEDKLGIRGNSCTVHF
NKTIPVENILGVDGDKVAMNINLGRFSMGSVVAGLLKRLIEMTAETACTRQPNK
RLSEGLIOEKFAKMAOKAVYVESMTYLTAGMLDQPPDCSIEAAMVKKVSSRAWQ
CVSEALOILGSLGYTRDYPYERILDRILILIFETNEILMYIALTLQHQHAGILTT
RIHEKQKAVSTVMDTVGRURRDSLRITVDLGTGNHGVHPSLADSAKFEENTYCF
LNTFTLLRFQKTIEMBEQLVLRKRVANILINLYGMTAVLSRASRSIRIGLRNHDHVL
GARTVEAYLQNLFLSOLDKYAPENLDEQIKKVSQIILEKRAYICAHPLDRTC"
BASE COUNT 624 a 617 c 682 g 525 t
ORIGIN
Query Match 100.0%; Score 1863; DB 9; Length 2448;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGGCTCGGGCTCTTCTCTCGCACACACGCGCTCGGCTCGTGGCTCGCGGGGTCG 60
Db 66 ATGAGCGGCTCGGGCTCTTCTCTCGCACACACGCGCTCGGCTCGTGGCTCGCGGGGTCG 125
QY 61 GTGCTCTTACCGCGAACCGCGGCTACTGCGCACACACGCGGCTGTACGAGCTTTTCGCC 120
Db 126 GTGCTCTTACCGCGAACCGCGGCTACTGCGCACACACGCGGCTGTACGAGCTTTTCGCC 185
QY 121 AAAGAGCTTTTCTCTAGGCAAAATCAAGAAGAAAGAACTTTTCCCATTTCCAGAGTTCG 180
Db 186 AAAGAGCTTTTCTCTAGGCAAAATCAAGAAGAAAGAACTTTTCCCATTTCCAGAGTTCG 245
QY 181 CAAGATGAATTAATGAATCAATCACTAGTCTTCTGGACCCCGTGGAAAAATTTCTTCACTGAA 240
Db 246 CAAGATGAATTAATGAATCAATCACTAGTCTTCTGGACCCCGTGGAAAAATTTCTTCACTGAA 305
QY 241 GAGTGGACTCCCGAAAAATTTGACGAGGAGGAAAAATCCACAGATGAAGCTTTGGAGAAA 300
Db 306 GAGTGGACTCCCGAAAAATTTGACGAGGAGGAAAAATCCACAGATGAAGCTTTGGAGAAA 365
QY 301 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAAGAGAAATATGTTGGCTTGGGCTTC 360
Db 366 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAAGAGAAATATGTTGGCTTGGGCTTC 425
QY 361 TCCAAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCTACTGTG 420
Db 426 TCCAAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCTACTGTG 485
QY 421 ACCCTGGCAGCGCACCGAGGCTATTGGCTTCAAGGGATCATCTTGGCTGGCACTGAGGAG 480
Db 486 ACCCTGGCAGCGCACCGAGGCTATTGGCTTCAAGGGATCATCTTGGCTGGCACTGAGGAG 545
QY 481 CAGAAAGCCAAATACTTGGCTTAACTGGCGTCCGGGAGACATTTGCACGCTTCTGCCTC 540
Db 546 CAGAAAGCCAAATACTTGGCTTAACTGGCGTCCGGGAGACATTTGCACGCTTCTGCCTC 605
QY 541 ACGGAGCCAGCCAGTGGGAGCGCATCGAGCTCAATCCGGAGCAGACCCACATAAGTGA 600
|||||

```

Db	606	ACGGACCCAGCCAGTGGGAGCGATGCGAGCCTCAATCCGGAGCAGACCCACACTAAGTGAA	665
Qy	601	GACAAGAAGCACTACATCTCAATGGCTCCAAGGCTTGGATTACTAATGGAGGACTGGCC	660
Db	666	GACAAGAAGCACTACATCTCAATGGCTCCAAGGCTTGGATTACTAATGGAGGACTGGCC	725
Qy	661	AATATTTTACTGTGTTTCAAGAGACTGAGTCTGTTGATTCTTGATGGATCACTGAAAGAC	720
Db	726	AATATTTTACTGTGTTTCAAGAGACTGAGTCTGTTGATTCTTGATGGATCACTGAAAGAC	785
Qy	721	AAATACACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAA	780
Db	786	AAATACACAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCCGAA	845
Qy	781	GATAAATTAGGCATTCGGGCTCCAACTGTGGAAGTCCATTTTCAAAACACCCAGATA	840
Db	846	GATAAATTAGGCATTCGGGCTCCAACTGTGGAAGTCCATTTTCAAAACACCCAGATA	905
Qy	841	CTGTGGAAGAACCTCTTGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC	900
Db	906	CTGTGGAAGAACCTCTTGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC	965
Qy	901	AACAGCGCCGGTTTCAGCATGGCAGCGTCGTGGCTGCTCAAGAGATTGATGAA	960
Db	966	AACAGCGCCGGTTTCAGCATGGCAGCGTCGTGGCTGCTCAAGAGATTGATGAA	1025
Qy	961	ATGACTGCTGAGTACGCTGCACAAGAAACAGTTTAAAGAGGCTCAGTGAATTGGA	1020
Db	1026	ATGACTGCTGAGTACGCTGCACAAGAAACAGTTTAAAGAGGCTCAGTGAATTGGA	1085
Qy	1021	TTGATTTCAGGAGAAATTCGACATGATGGCTCAGAAGGCTTACGTCATGAGAGATGACC	1080
Db	1086	TTGATTTCAGGAGAAATTCGACATGATGGCTCAGAAGGCTTACGTCATGAGAGATGACC	1145
Qy	1081	TACCTCAGCAGAGGATGCTGGACCAACCTGCTTCCGACATGCTCCATCGAGGAGCC	1140
Db	1146	TACCTCAGCAGAGGATGCTGGACCAACCTGCTTCCGACATGCTCCATCGAGGAGCC	1205
Qy	1141	ATGTGAAGGTGTTTCAGCTCCGAGCGCCCTGCGAGTGTGTGAGTGAGGCGCTGCAGATC	1200
Db	1206	ATGTGAAGGTGTTTCAGCTCCGAGCGCCCTGCGAGTGTGTGAGTGAGGCGCTGCAGATC	1265
Qy	1201	CTCGGGGCTTGGGCTACACAAGGAGATATCCGTACAGCGCATACTGCTGACACCCG	1260
Db	1266	CTCGGGGCTTGGGCTACACAAGGAGATATCCGTACAGCGCATACTGCTGACACCCG	1325
Qy	1261	ATCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGT	1320
Db	1326	ATCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGT	1385
Qy	1321	CTGCAGCATCGCGCGGATCCTGACTACCAAGATCATGAGCTTAAACAGGCGCAAGTG	1380
Db	1386	CTGCAGCATCGCGCGGATCCTGACTACCAAGATCATGAGCTTAAACAGGCGCAAGTG	1445
Qy	1381	AGCAGCTCATGATACGTTGGCGGAGGCTTCGGGACTCCCTGGCGCAACTGTGGAC	1440
Db	1446	AGCAGCTCATGATACGTTGGCGGAGGCTTCGGGACTCCCTGGCGCAACTGTGGAC	1505
Qy	1441	CTGGGGCTGACAGGCAACCATGAGTTGTGCACCCCACTGTGCGGACAGTGCACCAAG	1500
Db	1506	CTGGGGCTGACAGGCAACCATGAGTTGTGCACCCCACTGTGCGGACAGTGCACCAAG	1565
Qy	1501	TTTGAGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACATCTGCTCCGCTTTGGC	1560
Db	1566	TTTGAGGAGAACACCTACTGCTTCGGCGGACCGTGGAGACATCTGCTCCGCTTTGGC	1625
Qy	1561	AAGACCATCATGAGGAGCAGCTGGTACTGAAGGGGTGGCGCAACATCCTCATCAACCTG	1620
Db	1626	AAGACCATCATGAGGAGCAGCTGGTACTGAAGGGGTGGCGCAACATCCTCATCAACCTG	1685
Qy	1621	TATGCGATGACGGCGTGTGTCGGGGCGCAGCGCTCCATCCGCAITGGGCTCGGCAAC	1680
Db	1686	TATGCGATGACGGCGTGTGTCGGGGCGCAGCGCTCCATCCGCAITGGGCTCGGCAAC	1745

Qy	1681	CACGACCACAGGTTCTCTTGGCCAAACACCTTCTGCTGGAAGCTTACTTGCAGAACTC	1740
Db	1746	CACGACCACAGGTTCTCTTGGCCAAACACCTTCTGCTGGAAGCTTACTTGCAGAACTC	1805
Qy	1741	TTCAGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAAGAA	1800
Db	1806	TTCAGCCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAAGAA	1865
Qy	1801	GTGTCCCAAGCAGATCTTGAAGAGGAGCCTATATCTGTGCCACCTCTCTGCACAGGACA	1860
Db	1866	GTGTCCCAAGCAGATCTTGAAGAGGAGCCTATATCTGTGCCACCTCTCTGCACAGGACA	1925
Qy	1861	TGC 1863	
Db	1926	TGC 1928	

RESULT 3

AX394747

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

AX394747

Sequence 1 from Patent WO0218582.

AX394747

AX394747.1

GI:21065826

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1

Meyers, R. and Hunter, J.J.

62112, a novel human dehydrogenase and uses thereof

Patent: WO 0218582-A 1 07-MAR-2002;

Millennium Pharmaceuticals, Inc. (US)

Location/Qualifiers

1..2452

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

67..1932

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD32049.1"

/db_xref="GI:21065827"

/translation="MSGGGLFRTTAAARACRLVSVSTANRLRLTSPVRAFAKELF

LKIKKKEVPEPVSODELNEINFLGPVERFFTEVDSRKIDQEGKIPDETLEKLE

SUGFLQVPEYGGGLGFSNTMYSLGEIISMDGSTTTLAAHQGLKGIILAGTTE

QKAYLPKLASGEHIAAFLCTEPASGDAASRSRATLSEDKKHVILNGSKVMTNGG

LANIFTVFARTEVDSGSKDKITAFIVERDFGGVTNGKPEDKLGIRGSNTCEHFE

NKIPVENTILGEVGDGFKVAMNINLSRFSMSVAGLLRLIEMTAETACTRKQFNK

RSEFGLIQEKALMAQKAYVMSMTYLTAGMLDQDFPDCSIEAMVKVFSSEAAWQ

CYSEALIQGLGYTRDYPYERILTRILLIPETNEILRMVIALTGLOHAGRILT

RHELKQKSVTMDTVGRRLRDSIGRTVDLGTGNHGVVHPSLADSNKPFENTYCF

GRVETLLRFKGTIMEEQILVKRVANILNLYGMTAVILEKVSQIILEKRAYICAHPLDRT

LANTFCEAYLQNLFSLSQLDKYAPENLDEQIKVSSQIILEKRAYICAHPLDRT

BASE COUNT

616 a 616 c 685 g 535 t

ORIGIN

Query Match

100.0%; Score 1863; DB 6; Length 2452;

Best Local Similarity

100.0%; Pred No. 0;

Matches 1863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1

ATGAGCGGCTCGCGGCTCTTCTTCGCGCACACCGGCTCGGCTCGCTCGCGGGGTCTG

60

Db

67

ATGAGCGGCTCGCGGCTCTTCTTCGCGCACACCGGCTCGGCTCGCTCGCGGGGTCTG

136

Qy

61

GTGGTCTCTACCGGACCGCGGCTACTCGCACACCGGCTGTGTACGAGCTTTCGCC

120

Db

127

GTGGTCTCTACCGGACCGCGGCTACTCGCACACCGGCTGTGTACGAGCTTTCGCC

186

Qy

121

AAAGAGCTTTTCTTAGCAAAAATCAAGAAGAAAGAGTTCCTCCATTTCCAGAGTAGC

180

Db

187

AAAGAGCTTTTCTTAGCAAAAATCAAGAAGAAAGAGTTCCTCCATTTCCAGAGTAGC

246

181 CAAGATGAACCTTAATGAATCAATCAGTCTCTGGACCCGTGGAATAATCTTCACTGAA 240
182 |
183 |
184 |
185 |
186 |
187 |
188 |
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 |
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 |
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 GAGGTGACCTCCGGAATAATGACAGGAGGAAATCCAGATGAAATTTGGAGAAA 300
242 |
243 |
244 |
245 |
246 |
247 CAAGATGAACCTTAATGAATCAATCAGTCTCTGGACCCGTGGAATAATCTTCACTGAA 306
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 TTCAAGAGCCTAGGCTTTTGGGCTTGAAGTCCCAAGAGATATGGTGGCTGGGCTTC 360
302 |
303 |
304 |
305 |
306 |
307 GAGGTGACCTCCGGAATAATGACAGGAGGAAATCCAGATGAAATTTGGAGAAA 366
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 TCCAACACCATGATCAAGACTAGGGAGATCATCAGCATGGATGGTGGCTTCACTGTC 420
362 |
363 |
364 |
365 |
366 |
367 TTCAAGAGCCTAGGCTTTTGGGCTTGAAGTCCCAAGAGATATGGTGGCTGGGCTTC 426
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 ACCCTGGCAGCCAGCAGGCTATTTGGCTTCAAGGGATCATCTTGGCTGGCCTGAGGAG 480
422 |
423 |
424 |
425 |
426 |
427 TCCAACACCATGATCAAGACTAGGGAGATCATCAGCATGGATGGTGGCTTCACTGTC 486
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 CAGAAAGCCAAATACCTTAACTGAGCTTCCGGGAGCAGATTCAGTCCGAGGAGGAGG 540
482 |
483 |
484 |
485 |
486 |
487 ACCCTGGCAGCCAGCAGGCTATTTGGCTTCAAGGGATCATCTTGGCTGGCCTGAGGAG 546
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 ACGGAGCCAGCAGTGGGAGGATGAGCTTCAATCCGAGGAGGAGGAGGAGGAGGAGG 600
542 |
543 |
544 |
545 |
546 |
547 CAGAAAGCCAAATACCTTAACTGAGCTTCCGGGAGCAGATTCAGTCCGAGGAGGAGG 606
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 GACAAGACCACTACATCTCAATGGCTTCCAGGCTTGGATTAATGGAGGAGTGGCC 660
602 |
603 |
604 |
605 |
606 |
607 ACGGAGCCAGCAGTGGGAGGATGAGCTTCAATCCGAGGAGGAGGAGGAGGAGGAGG 666
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 AATATTTTACTGTGTTGCAAGACTGAGGTGCTTGAATTCATGATGGATCACTGAAAGAC 720
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
688 |
689 |
690 |
691 |
692 |
693 |
694 |
695 |
696 |
697 |
698 |
699 |
700 |
701 |
702 |
703 |
704 |
705 |
706 |
707 |
708 |
709 |
710 |
711 |
712 |
713 |
714 |
715 |
716 |
717 |
718 |
719 |
720 |
721 AAATACAGACATTCATAGTAGAAGACATTTGGTGGAGTCACTTAATGGGAACCCGAA 780
722 |
723 |
724 |
725 |
726 |
727 AATATTTTACTGTGTTGCAAGACTGAGGTGCTTGAATTCATGATGGATCACTGAAAGAC 786
728 |
729 |
730 |
731 |
732 |
733 |
734 |
735 |
736 |
737 |
738 |
739 |
740 |
741 |
742 |
743 |
744 |
745 |
746 |
747 |
748 |
749 |
750 |
751 |
752 |
753 |
754 |
755 |
756 |
757 |
758 |
759 |
760 |
761 |
762 |
763 |
764 |
765 |
766 |
767 |
768 |
769 |
770 |
771 |
772 |
773 |
774 |
775 |
776 |
777 |
778 |
779 |
780 |
781 GATAAATAGGCATTCGGGCTCCCAACACTTGTGAAGTCCATTTTGAACACCAAGATA 840
782 |
783 |
784 |
785 |
786 |
787 AAATACAGACATTCATAGTAGAAGACATTTGGTGGAGTCACTTAATGGGAACCCGAA 846
788 |
789 |
790 |
791 |
792 |
793 |
794 |
795 |
796 |
797 |
798 |
799 |
800 |
801 |
802 |
803 |
804 |
805 |
806 |
807 |
808 |
809 |
810 |
811 |
812 |
813 |
814 |
815 |
816 |
817 |
818 |
819 |
820 |
821 |
822 |
823 |
824 |
825 |
826 |
827 |
828 |
829 |
830 |
831 |
832 |
833 |
834 |
835 |
836 |
837 |
838 |
839 |
840 |
841 CCTGTGGAACATCCTTTGGAGAGCTCGGAGATGGGTTTAAAGTGGCCATGAACATCTC 900
842 |
843 |
844 |
845 |
846 |
847 GATAAATAGGCATTCGGGCTCCCAACACTTGTGAAGTCCATTTTGAACACCAAGATA 906
848 |
849 |
850 |
851 |
852 |
853 |
854 |
855 |
856 |
857 |
858 |
859 |
860 |
861 |
862 |
863 |
864 |
865 |
866 |
867 |
868 |
869 |
870 |
871 |
872 |
873 |
874 |
875 |
876 |
877 |
878 |
879 |
880 |
881 |
882 |
883 |
884 |
885 |
886 |
887 |
888 |
889 |
890 |
891 |
892 |
893 |
894 |
895 |
896 |
897 |
898 |
899 |
900 |
901 AACACGGCCGCTTCAGCATGGCAGCGCTGCTGGCTGGGCTGCTCAAGAGATTTGAA 960
902 |
903 |
904 |
905 |
906 |
907 CCTGTGGAACATCCTTTGGAGAGCTCGGAGATGGGTTTAAAGTGGCCATGAACATCTC 966
908 |
909 |
910 |
911 |
912 |
913 |
914 |
915 |
916 |
917 |
918 |
919 |
920 |
921 |
922 |
923 |
924 |
925 |
926 |
927 |
928 |
929 |
930 |
931 |
932 |
933 |
934 |
935 |
936 |
937 |
938 |
939 |
940 |
941 |
942 |
943 |
944 |
945 |
946 |
947 |
948 |
949 |
950 |
951 |
952 |
953 |
954 |
955 |
956 |
957 |
958 |
959 |
960 |
961 ATGACTGCTGAGTACGCTGCAAGAGAAACAGTTTAAAGAGGCTCAGTGAATTTGGA 1020
962 |
963 |
964 |
965 |
966 |
967 AACACGGCCGCTTCAGCATGGCAGCGCTGCTGGCTGGGCTGCTCAAGAGATTTGAA 1026
968 |
969 |
970 |
971 |
972 |
973 |
974 |
975 |
976 |
977 |
978 |
979 |
980 |
981 |
982 |
983 |
984 |
985 |
986 |
987 |
988 |
989 |
990 |
991 |
992 |
993 |
994 |
995 |
996 |
997 |
998 |
999 |
1000 |
1001 |
1002 |
1003 |
1004 |
1005 |
1006 |
1007 |
1008 |
1009 |
1010 |
1011 |
1012 |
1013 |
1014 |
1015 |
1016 |
1017 |
1018 |
1019 |
1020 |
1021 TTGATTCAGGAGAAATTTGCACTGATGCTCAGAGGCTTACAGAGGCTCAGTGAATTTGGA 1086
1022 |
1023 |
1024 |
1025 |
1026 |
1027 ATGACTGCTGAGTACGCTGCAAGAGAAACAGTTTAAAGAGGCTCAGTGAATTTGGA 1086
1028 |
1029 |
1030 |
1031 |
1032 |
1033 |
1034 |
1035 |
1036 |
1037 |
1038 |
1039 |
1040 |
1041 |
1042 |
1043 |
1044 |
1045 |
1046 |
1047 |
1048 |
1049 |
1050 |
1051 |
1052 |
1053 |
1054 |
1055 |
1056 |
1057 |
1058 |
1059 |
1060 |
1061 |
1062 |
1063 |
1064 |
1065 |
1066 |
1067 |
1068 |
1069 |
1070 |
1071 |
1072 |
1073 |
1074 |
1075 |
1076 |
1077 |
1078 |
1079 |
1080 |
1081 TACCTCAGCAGGAGATGCTGACCAACCTTCCGACTGCTCCATCGAGGAGCC 1140
1082 |
1083 |
1084 |
1085 |
1086 |
1087 TTGATTCAGGAGAAATTTGCACTGATGCTCAGAGGCTTACAGAGGCTCAGTGAATTTGGA 1146
1088 |
1089 |
1090 |
1091 |
1092 |
1093 |
1094 |
1095 |
1096 |
1097 |
1098 |
1099 |
1100 |
1101 |
1102 |
1103 |
1104 |
1105 |
1106 |
1107 |
1108 |
1109 |
1110 |
1111 |
1112 |
1113 |
1114 |
1115 |
1116 |
1117 |
1118 |
1119 |
1120 |
1121 |
1122 |
1123 |
1124 |
1125 |
1126 |
1127 |
1128 |
1129 |
1130 |
1131 |
1132 |
1133 |
1134 |
1135 |
1136 |
1137 |
1138 |
1139 |
1140 |
1141 ATGTTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGGCGCTGCAGATC 1200
1142 |
1143 |
1144 |
1145 |
1146 |
1147 TACCTCAGCAGGAGATGCTGGACCAACCTTCCGACTGCTCCATCGAGGAGCC 1206
1148 |
1149 |
1150 |
1151 |
1152 |
1153 |
1154 |
1155 |
1156 |
1157 |
1158 |
1159 |
1160 |
1161 |
1162 |
1163 |
1164 |
1165 |
1166 |
1167 |
1168 |
1169 |
1170 |
1171 |
1172 |
1173 |
1174 |
1175 |
1176 |
1177 |
1178 |
1179 |
1180 |
1181 |
1182 |
1183 |
1184 |
1185 |
1186 |
1187 |
1188 |
1189 |
1190 |
1191 |
1192 |
1193 |
1194 |
1195 |
1196 |
1197 |
1198 |
1199 |
1200 |
1201 CTCCGGGCGCTTGGCTACACAGGAGGACTATCCGTACAGGCGCATACTGCGGTGACACCCG 1260
1202 |
1203 |
1204 |
1205 |
1206 |
1207 ATGTTGAAGTGTTCAGCTCCGAGGCGCCCTGGCAGTGTGTGAGTGAGGCGCTGCAGATC 1266
1208 |
1209 |
1210 |
1211 |
1212 |
1213 |
1214 |
1215 |
1216 |
1217 |
1218 |
1219 |
1220 |
1221 |
1222 |
1223 |
1224 |
1225 |
1226 |
1227 |
1228 |
1229 |
1230 |
1231 |
1232 |
1233 |
1234 |
1235 |
1236 |
1237 |
1238 |
1239 |
1240 |
1241 |
1242 |
1243 |
1244 |
1245 |
1246 |
1247 |
1248 |
1249 |
1250 |
1251 |
1252 |
1253 |
1254 |
1255 |
1256 |
1257 |
1258 |
1259 |
1260 |
1261 |
1262 |
1263 |
1264 |
1265 |
1266 |
1267 CTCCGGGCGCTTGGCTACACAGGAGGACTATCCGTACAGGCGCATACTGCGGTGACACCCG 1326
1268 |
1269 |
1270 |
1271 |
1272 |
1273 |
1274 |
1275 |
1276 |
1277 |
1278 |
1279 |
1280 |
1281 |
1282 |
1283 |
1284 |
1285 |
1286 |
1287 |
1288 |
1289 |
1290 |
1291 |
1292 |
1293 |
1294 |
1295 |
1296 |
1297 |
1298 |
1299 |
1300 |
1301 |
1302 |
1303 |
1304 |
1305 |
1306 |
1307 |
1308 |
1309 |
1310 |
1311 |
1312 |
1313 |
1314 |
1315 |
1316 |
1317 |
1318 |
1319 |
1320 |
1321 |
1322 |
1323 |
1324 |
1325 |
1326 |
1327 |
1328 |
1329 |
1330 |
1331 |
1332 |
1333 |
1334 |
1335 |
1336 |
1337 |
1338 |
1339 |
1340 |
1341 |
1342 |
1343 |
1344 |
1345 |
1346 |
1347 |
1348 |
1349 |
1350 |
1351 |
1352 |
1353 |
1354 |
1355 |
1356 |
1357 |
1358 |
1359 |
1360 |
1361 |
1362 |
1363 |
1364 |
1365 |
1366 |
1367 |
1368 |
1369 |
1370 |
1371 |
1372 |
1373 |
1374 |
1375 |
1376 |
1377 |
1378 |
1379 |
1380 |
1381 |
1382 |
1383 |
1384 |
1385 |
1386 |
1387 |
1388 |
1389 |
1390 |
1391 |
1392 |
1393 |
1394 |
1395 |
1396 |
1397 |
1398 |
1399 |
1400 |
1401 |
1402 |
1403 |
1404 |
1405 |
1406 |
1407 |
1408 |
1409 |
1410 |
1411 |
1412 |
1413 |
1414 |
1415 |
1416 |
1417 |
1418 |
1419 |
1420 |
1421 |
1422 |
1423 |
1424 |
1425 |
1426 |
1427 |
1428 |
1429 |
1430 |
1431 |
1432 |
1433 |
1434 |
1435 |
1436 |
1437 |
1438 |
1439 |
1440 |
1441 |
1442 |
1443 |
1444 |
1445 |
1446 |
1447 |
1448 |
1449 |
1450 |
1451 |
1452 |
1453 |
1454 |
1455 |
1456 |
1457 |
1458 |
1459 |
1460 |
1461 |
1462 |
1463 |
1464 |
1465 |
1466 |
1467 |
1468 |
1469 |
1470 |
1471 |
1472 |
1473 |
1474 |
1475 |
1476 |
1477 |
1478 |
1479 |
1480 |
1481 |
1482 |
1483 |
1484 |
1485 |
1486 |
1487 |
1488 |
1489 |
1490 |
1491 |
1492 |
1493 |
1494 |
1495 |
1496 |
1497 |
1498 |
1499 |
1500 |
1501 |
1502 |
1503 |
1504 |
1505 |
1506 |
1507 |
1508 |
1509 |
1510 |
1511 |
1512 |
1513 |
1514 |
1515 |
1516 |
1517 |
1518 |
1519 |
1520 |
1521 |
1522 |
1523 |
1524 |
1525 |
1526 |
1527 |
1528 |
1529 |
1530 |
1531 |
1532 |
1533 |
1534 |
1535 |
1536 |
1537 |
1538 |
1539 |
1540 |
1541 |
1542 |
1543 |
1544 |
1545 |
1546 |
1547 |
1548 |
1549 |
1550 |
1551 |
1552 |
1553 |
1554 |
1555 |
1556 |
1557 |
1558 |
1559 |
1560 |
1561 |
1562 |
1563 |
1564 |
1565 |
1566 |
1567 |
1568 |
1569 |
1570 |
1571 |
1572 |
1573 |
1574 |
1575 |
1576 |
1577 |
1578 |
1579 |
1580 |
1581 |
1582 |
1583 |
1584 |
1585 |
1586 |
1587 |
1588 |
1589 |
1590 |
1591 |
1592 |
1593 |
1594 |
1595 |
1596 |
1597 |
1598 |
1599 |
1600 |
1601 |
1602 |
1603 |
1604 |
1605 |
1606 |
1607 |
1608 |
1609 |
1610 |
1611 |
1612 |
1613 |
1614 |
1615 |
1616 |
1617 |
1618 |
1619 |
1620 |
1621 |
1622 |
1623 |
1624 |
1625 |
1626 |
1627 |
1628 |
1629 |
1630 |
1631 |
1632 |
1633 |
1634 |
1635 |
1636 |
1637 |
1638 |
1639 |
1640 |
1641 |
1642 |
1643 |
1644 |
1645 |
1646 |
1647 |
1648 |
1649 |
1650 |
1651 |
1652 |
1653 |
1654 |
1655 |
1656 |
1657 |
1658 |
1659 |
1660 |
1661 |
1662 |<

PC C12P21/02,C1201/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers

FEATURES
source 1..2399
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 597 a 606 c 672 g 524 t
ORIGIN

Query Match	99.8%;	Score 1859.8;	DB 6;	Length 2399;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1861;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	ATGAGCGGCTGGGGCTCTTCTCGCCACACCGGCTGCGGCTCGTGCCTGCGGGGTCTG	60	
DB	47	ATGAGCGGCTGGGGCTCTTCTCGCCACACCGGCTGCGGCTCGTGCCTGCGGGGTCTG	106	
QY	61	GTGGTCTCTACCGCGAACCGGGCTACTGCGCACCAGCCCGCCGTAGAGCTTCGCC	120	
DB	107	GTGGTCTCTACCGCGAACCGGGCTACTGCGCACCAGCCCGCCGTAGAGCTTCGCC	166	
QY	121	AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAAGAGTTTTCCTATTTCCAGAAAGTTAGC	180	
DB	167	AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAAGAGTTTTCCTATTTCCAGAAAGTTAGC	226	
QY	181	CAAGATGAACCTTAATGAATCAATAGTTCTTGGGACCCGCTGGGAAATTTCTCACTGAA	240	
DB	227	CAAGATGAACCTTAATGAATCAATAGTTCTTGGGACCCGCTGGGAAATTTCTCACTGAA	286	
QY	241	GAGTGGACTCCGAAATTTGACAGAGGAGGAAATCCAGATGAAATTTGGAGAAA	300	
DB	287	GAGTGGACTCCGAAATTTGACAGAGGAGGAAATCCAGATGAAATTTGGAGAAA	346	
QY	301	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAGAGAAATATGGTGGCCTGGGCTTC	360	
DB	347	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAGAGAAATATGGTGGCCTGGGCTTC	406	
QY	361	TCCAACACCATGTACTCAAGCTAGGGAGATCATCAGCATGGATGGTCCATCACTGTG	420	
DB	407	TCCAACACCATGTACTCAAGCTAGGGAGATCATCAGCATGGATGGTCCATCACTGTG	466	
QY	421	ACCTTGCAGCGCACCAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCACTGAGGAG	480	
DB	467	ACCTTGCAGCGCACCAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCACTGAGGAG	526	
QY	481	CAGAAAGCCAAATACTTGGCTAACTGGCGTCCGGGAGCACATTCAGGCTTCTGCTC	540	
DB	527	CAGAAAGCCAAATACTTGGCTAACTGGCGTCCGGGAGCACATTCAGGCTTCTGCTC	586	
QY	541	ACGAGCGCAGCGAGTGGGCGATGACGCTCAATCCGAGCAGAGCCACACTAAAGTAA	600	
DB	587	ACGAGCGCAGCGAGTGGGCGATGACGCTCAATCCGAGCAGAGCCACACTAAAGTAA	646	
QY	601	GACAAGAGCACTACATCTCAATGGCTCCAAAGTCTGGATTTACTATGAGAGACTGGCC	660	
DB	647	GACAAGAGCACTACATCTCAATGGCTCCAAAGTCTGGATTTACTATGAGAGACTGGCC	706	
QY	661	AATATTTTACTGTGTTGCAAGACATGAGTCTGATTTCTGATGATGATGATGATGATG	720	
DB	707	AATATTTTACTGTGTTGCAAGACATGAGTCTGATTTCTGATGATGATGATGATGATG	766	
QY	721	AAATATCAGCATTTCAATAGTAGAAGAGACTTTGGTGGAGTCTACTAATGGAACCCGAA	780	
DB	767	AAATATCAGCATTTCAATAGTAGAAGAGACTTTGGTGGAGTCTACTAATGGAACCCGAA	826	
QY	781	GATAAATTAGGCAATTCGGGCTCCAAACACTTGTGAAGTCCATTTGAAACACCAAGATA	840	
DB	827	GATAAATTAGGCAATTCGGGCTCCAAACACTTGTGAAGTCCATTTGAAACACCAAGATA	886	
QY	841	CCTGTGGAACATCTTGGAGAGGTGGAGATGGGTTTAAGGTGGCCATGAACATCCTC	900	

RESULT 5

DB	887	CCTGTGGAACATCTTGGAGAGGTGCGAGATGGTTAAGTGGCCATGAACATCCTC	946	
QY	901	AACAGCGCCGGTTTCAGCATGGCAGCGTCTGCTGGCTGCTCAAGAGATTGATTGAA	960	
DB	947	AACAGCGCCGGTTTCAGCATGGCAGCGTCTGCTGGCTGCTCAAGAGATTGATTGAA	1006	
QY	961	ATGACTGCTGAGTACGCTGCACAGAGAAACAGTTTAAACAAGAGGCTCAGTGAATTTGGA	1020	
DB	1007	ATGACTGCTGAGTACGCTGCACAGAGAAACAGTTTAAACAAGAGGCTCAGTGAATTTGGA	1066	
QY	1021	TTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTTCATGAGAGTATGACC	1080	
DB	1067	TTGATTTCAGGAGAAATTTGCACTGATGGCTCAGAGGCTTACGTTCATGAGAGTATGACC	1126	
QY	1081	TACCTTCACAGCAGGATGCTTGACCAACCTGCTTCCGACTGCTCCATCGAGCAGCC	1140	
DB	1127	TACCTTCACAGCAGGATGCTTGACCAACCTGCTTCCGACTGCTCCATCGAGCAGCC	1186	
QY	1141	ATGSGTGAAGTGTTCAGCTCCGAGCCGCTGGCAGTGTGTGAGTGAGCGCTGCAGATC	1200	
DB	1187	ATGSGTGAAGTGTTCAGCTCCGAGCCGCTGGCAGTGTGTGAGTGAGCGCTGCAGATC	1246	
QY	1201	CTCGGGGCTTGGGCTACACAGGAGCTATCGTACGAGCGCATACTGCTGACACCCGC	1260	
DB	1247	CTCGGGGCTTGGGCTACACAGGAGCTATCGTACGAGCGCATACTGCTGACACCCGC	1306	
QY	1261	ATCCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGAGTGTACATCCCTGACGGGT	1320	
DB	1307	ATCCTCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGAGTGTACATCCCTGACGGGT	1366	
QY	1321	CTGAGCATGCGCGCGCATCTCTGACTACCAGATCCATGAGCTTAAACAGGCCAAATG	1380	
DB	1367	CTGAGCATGCGCGCGCATCTCTGACTACCAGATCCATGAGCTTAAACAGGCCAAATG	1426	
QY	1381	AGCAGCATCATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGCGGAACTGTGGAC	1440	
DB	1427	AGCAGCATCATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGCGGAACTGTGGAC	1486	
QY	1441	CTGGGCTCAGAGGCAACCATGAGTTGTGACCCCGAGTCTTGGGAGCATGCCAACAG	1500	
DB	1487	CTGGGCTCAGAGGCAACCATGAGTTGTGACCCCGAGTCTTGGGAGCATGCCAACAG	1546	
QY	1501	TTTGAGGAGAACACCTTACTGCTTCGGCCGAGCCGTGGAGACACTGCTGCTCGCTTTGGC	1560	
DB	1547	TTTGAGGAGAACACCTTACTGCTTCGGCCGAGCCGTGGAGACACTGCTGCTCGCTTTGGC	1606	
QY	1561	AAGACCATCATGGAGGAGCAGCTGTGTAAGCGGGTGGCAACATCTCATCAACCTG	1620	
DB	1607	AAGACCATCATGGAGGAGCAGCTGTGTAAGCGGGTGGCAACATCTCATCAACCTG	1666	
QY	1621	TATGGCATGACGGCCGTGCTGCGCGGCGCAGCCGCTCCATCCGATTTGGGCTCCGCAAC	1680	
DB	1667	TATGGCATGACGGCCGTGCTGCGCGGCGCAGCCGCTCCATCCGATTTGGGCTCCGCAAC	1726	
QY	1681	CAGCAGCAGAGGTTCTTGGCCAAACACCTTCTGCTGGAAGCTTACTTTGAGAAATCTC	1740	
DB	1727	CAGCAGCAGAGGTTCTTGGCCAAACACCTTCTGCTGGAAGCTTACTTTGAGAAATCTC	1786	
QY	1741	TTGAGGCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAGAAA	1800	
DB	1787	TTGAGGCTCTCTCAGCTGGCAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAGAAA	1846	
QY	1801	GTGTCACAGCAGATCTTGAAGAAGCGAGCCTATATCTGTGCCACCCCTCTGGACAGACA	1860	
DB	1847	GTGTCACAGCAGATCTTGAAGAAGCGAGCCTATATCTGTGCCACCCCTCTGGACAGACA	1906	
QY	1861	TGC 1863		
DB	1907	TGC 1909		

AK024012	LOCUS	AK024012	2399 bp	mRNA	linear	PRI 01-AUG-2002
DEFINITION		Human sapiens cDNA FLJ13950 fls, clone Y79AA1001048, weakly similar to ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC PRECURSOR (EC 1.3.99.-)				
ACCESSION		AK024012	GI:10436257			
KEYWORDS		oligo capping; fls (full insert sequence)				
SOURCE		Human sapiens				
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE		1	Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoliri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T. NEDO human cDNA sequencing project			
AUTHORS		Unpublished				
REFERENCE		2 (bases 1 to 2399)				
AUTHORS		Isogai, T. and Otsuki, T.				
TITLE		Direct Submission				
JOURNAL		Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)				
COMMENT		NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.				
FEATURES		Location/Qualifiers				
source		1..2399				
		/organism="Homo sapiens"				
		/mol_type="mRNA"				
		/db_xref="taxon:9606"				
		/clone="Y79AA1001048"				
		/cell_line="Y79"				
		/cell_type="retinoblastoma"				
		/clone_lib="Y79AAL"				
		/note="cloning vector: pME18SFL3"				
		47..1912				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="BAB14775.1"				
		/db_xref="GI:10436258"				
		/translation="MSGGGLFLRTTAAARCGLVVSTANRLRLTSPVRAFAKELF LKIKKEVFPPEVSDLEINQFLGPVEKFFTEEDSRKIDQEKIPDLEKLEK SLGLFQVPEYGGSGFNTMYSLRGEIISWDGSIITVLAHQAIQGLKGIILAGTEE QAKYLPKLESHLEAFLTEPASGDAASIRSRATLSEDKHYILNGSKVITNGG LANITVFATKEVVDGSKVDKITAFTVERDQFGVNGKPEDKLGRGNTVEHFE NTKIPVENLGEVGDGFKVAMNLSGRFSVMSVYAGLLKRLIEATYATCRKOFNK RLSERGLQEFALMAQRAYVHESMYTLTGMDDQPPDCIEAAMVYFSSAAWQ CVSAQLTGLGGVTRIDYRILRDLTRILFETNELRMVIALTQLQHAGRLIT RHLEKQAKVTDVMDTRRLRSLGRVTLGLTGNHGVHPHSLADSANKPEENTYCF GRVETLLLRFGKTIEMEQVLKRVANILINLYGMTAVLSRSRISIRGLRNHDEVL LANITFCVEAYLQNLFSLSOLDKYAPENLDEIKKVSQILEKRAYICAHPLDRVC"				
BASE COUNT		597 a 606 c 672 g 524 t				
ORIGIN						
Query Match		99.8%; Score 1859.8; DB 9; Length 2399;				
Best Local Similarity		99.9%; Pred. No. 0;				
Matches 1861; Conservative		0; Mismatches 2; Indels 0; Gaps 0;				
Qy		1 ATGAGCGGCTCGGGCTCTTCTCGGCACACAGCGCTGCGGCTGCTGCTGCGCGGGGTCTG 60				
Db						
		47 ATGAGCGGCTCGGGCTCTTCTCGGCACACAGCGCTGCGGCTGCTGCTGCGCGGGGTCTG 106				
Qy		61 GTGCTCTCTACCGCGGAACCGCGGCTACTTGGCCACACAGCGCGCTGTACGAGCTTTTCGCC 120				
Db		107 GTGCTCTCTACCGCGGAACCGCGGCTACTTGGCCACACAGCGCGCTGTACGAGCTTTTCGCC 166				
Qy		121 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGAGTTTTCCTCATTTCCAGAGTTAGC 180				
Db						
		167 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGAGTTTTCCTCATTTCCAGAGTTAGC 226				
Qy		181 CAAGATGAATTAAGTAAATCAATCAAGTTCTTGGGACCGCTGGGAAATATCTTCACTGAA 240				
Db						
		227 CAAGATGAATTAAGTAAATCAATCAAGTTCTTGGGACCGCTGGGAAATATCTTCACTGAA 286				
Qy		241 GAGGTGGACTCCCGAAATATGACACAGGAGGAAATCCAGATGCAAACTTTTGGAGAA 300				
Db						
		287 GAGGTGGACTCCCGAAATATGACACAGGAGGAAATCCAGATGCAAACTTTTGGAGAA 346				
Qy		301 TTGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTTGGCTTC 360				
Db						
		347 TTGAAGAGCTAGGCTTTTGGGCTGCAAGTCCAGAGAAATATGGTGGCTTGGCTTC 406				
Qy		361 TCCACACCATCTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACTGTG 420				
Db						
		407 TCCACACCATCTACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCATCACTGTG 466				
Qy		421 ACCCTGGCAGCCACAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG 480				
Db						
		467 ACCCTGGCAGCCACAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG 526				
Qy		481 CAGAAGCCAAATACTTGCCTAAACTGCGCTCGGGGAGCACATTTGCAGCCTTTCGCCTC 540				
Db						
		527 CAGAAGCCAAATACTTGCCTAAACTGCGCTCGGGGAGCACATTTGCAGCCTTTCGCCTC 586				
Qy		541 ACGGAGCAGCCAGTGGGAGCGATCCAGCCTCAATCCGAGGAGGAGGAGGAGGAGGAGG 600				
Db						
		587 ACGGAGCAGCCAGTGGGAGCGATCCAGCCTCAATCCGAGGAGGAGGAGGAGGAGGAGG 646				
Qy		601 GACAAAGACACTACATCTCTCAATGGCTCCAAAGGTCTGGATTACTTAATGGAGGACTGGCC 660				
Db						
		647 GACAAAGACACTACATCTCTCAATGGCTCCAAAGGTCTGGATTACTTAATGGAGGACTGGCC 706				
Qy		661 AATATTTTACTGTGTTTGCAGAACTGAGGTGTTGATCTGTATGATGAGTCAAGTCAAGAC 720				
Db						
		707 AATATTTTACTGTGTTTGCAGAACTGAGGTGTTGATCTGTATGATGAGTCAAGTCAAGAC 766				
Qy		721 AAAATACAGCATTCATAGTAGAAAGACACTTTTGTGGAGTCACTTAATGGGAAACCCGAA 780				
Db						
		767 AAAATACAGCATTCATAGTAGAAAGACACTTTTGTGGAGTCACTTAATGGGAAACCCGAA 826				
Qy		781 GATAAATPAGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGATA 840				
Db						
		827 GATAAATPAGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGATA 886				
Qy		841 CCTGTGGAACATCTTGGAGAGGTGGAGATCGGTTTAAAGTGGCCATGAACATCCTC 900				
Db						
		887 CCTGTGGAACATCTTGGAGAGGTGGAGATCGGTTTAAAGTGGCCATGAACATCCTC 946				
Qy		901 AACAGCGCCGGTTTTCAGCATGGCAGCGTCTGTGCTGGCTGCTCAAGAGATTCATTTGAA 960				
Db						
		947 AACAGCGCCGGTTTTCAGCATGGCAGCGTCTGTGCTGGCTGCTCAAGAGATTCATTTGAA 1006				
Qy		961 ATGACTGCTGAGTACGGCTTGCACAAAGAAACAGTTTAAACAGAGGCTCAGTGAATTTGA 1020				
Db						
		1007 ATGACTGCTGAGTACGGCTTGCACAAAGAAACAGTTTAAACAGAGGCTCAGTGAATTTGA 1066				
Qy		1021 TTGATTCAGGAGAAATTTGCATGCTGATGCTCAGAGGCTTACGTCATGGAGAGTATGACC 1080				
Db						
		1067 TTGATTCAGGAGAAATTTGCATGCTGATGCTCAGAGGCTTACGTCATGGAGAGTATGACC 1126				
Qy		1081 TACCTCAGCAGGAGTCTGGACCAACCTTTCCTCCGACCTGCTCATCGAGGAGCC 1140				
Db						
		1127 TACCTCAGCAGGAGTCTGGACCAACCTTTCCTCCGACCTGCTCATCGAGGAGCC 1186				
Qy		1141 ATGCTGAAGTGTTCAGCTCCGAGGCGCTTGGAGTGTGTGAGTGAAGGCGCTTCGACATC 1200				


```

Db      1187 ATGGTGAAGGCTTCAGCTCCGAGCGCGCTGGCAGTGTGTGAGTGAGCGCGCTGCAGATC 1246
QY      1201 CTCGGGGGCTTGGGCTACACAGGAGGACTATCCGTACGAGCGCATCTCGGTGACACCGC 1260
Db      1247 CTCGGGGGCTTGGGCTACACAGGAGGACTATCCGTACGAGCGCATCTCGGTGACACCGC 1306
QY      1261 ATCCCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGT 1320
Db      1307 ATCCCTCTCATCTTCGAGGGAACCAATGAGATTCTCCGGATGTACATCGCCCTGACGGGT 1366
QY      1321 CTGCAGATGCGCGGCGCATCTCTGACTACCGAGATCCATGAGCTTAACAGGCCAAAGTG 1380
Db      1367 CTGCAGATGCGCGGCGCATCTCTGACTACCGAGATCCATGAGCTTAACAGGCCAAAGTG 1426
QY      1381 AGCACAGTATGATACCTGCTTCGCGGAGGCTTCGGGACTCCCTGGCGGCAACTGTGGAC 1440
Db      1427 AGCACAGTATGATACCTGCTTCGCGGAGGCTTCGGGACTCCCTGGCGGCAACTGTGGAC 1486
QY      1441 CTGGGCTGACAGGCAACCATGAGTTGTGCACCCAGTCTTGGCGACAGTGCACCAAG 1500
Db      1487 CTGGGCTGACAGGCAACCATGAGTTGTGCACCCAGTCTTGGCGACAGTGCACCAAG 1546
QY      1501 TTGAGGAGAACACCTACTGCTTCGCGGAGGCTTCGGGACTCCCTGGCGGCAACTGTGGAC 1560
Db      1547 TTGAGGAGAACACCTACTGCTTCGCGGAGGCTTCGGGACTCCCTGGCGGCAACTGTGGAC 1606
QY      1561 AAGACCATCATGAGGAGCAGCTGTTGTAAGCGGGTGGCCAAACATCTCATCAACCTG 1620
Db      1607 AAGACCATCATGAGGAGCAGCTGTTGTAAGCGGGTGGCCAAACATCTCATCAACCTG 1666
QY      1621 TATGGCATGACGGCGCTGCTGCGGGGCGAGCGCTTCCATCGGATTTGGCTCCGCAAC 1680
Db      1667 TATGGCATGACGGCGCTGCTGCGGGGCGAGCGCTTCCATCGGATTTGGCTCCGCAAC 1726
QY      1681 CACGACACGAGGTTCTCTTGGCCAAACACTTCTGCGTGGAACTTACCTTGAGAACTCTC 1740
Db      1727 CACGACACGAGGTTCTCTTGGCCAAACACTTCTGCGTGGAACTTACCTTGAGAACTCTC 1786
QY      1741 TTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATTAAAGAA 1800
Db      1787 TTCAGGCTCTCTCAGCTGGACAAAGTATGCTCCAGAAACCTAGATGAGCAGATTAAAGAA 1846
QY      1801 GTCTCCACAGATCTCTGAGAGCGAGCTATATCTGTGCCACCTCTGGACAGGACA 1860
Db      1847 GTCTCCACAGATCTCTGAGAGCGAGCTATATCTGTGCCACCTCTGGACAGGACA 1906
QY      1861 TGC 1863
Db      1907 TGC 1909

RESULT 6
BC007970
LOCUS      2494 bp      mRNA      linear      PRI 20-MAY-2003
DEFINITION Homo sapiens very-long-chain acyl-CoA dehydrogenase VLCAD, mRNA
            (CDNA clone MGC:14452 IMAGE:4304209), complete cds.
ACCESSION BC007970
VERSION   BC007970.1 GI:14044101
KEYWORDS  MGC.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2494)
AUTHORS   Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
           Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
           Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
           Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
           Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
           Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
           Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
           Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
           Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

```

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, O., Small, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2494)

Strausberg, R.

Direct Submission

Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@hri.nih.gov

Akter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 20 Row: p Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

1. .2494

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14452 IMAGE:4304209"

/tissue_type="uterus, leiomyosarcoma"

/clone_lib="NIH_MGC_46"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1. .2494

/gene="VLCAD"

/note="Synonyms: NPD002, MGC14452, ACAD-9"

/db_xref="LocusID:28976"

89. .1954

/codon_start=1

/product="very-long-chain acyl-CoA dehydrogenase VLCAD"

/protein_id="AAH04790.1"

/db_xref="GI:14044102"

/db_xref="LocusID:28976"

/translation="MSGCGLFLRTAAARACRGLVSTANRLRLTSPVPYAFAKELK LGIKKKEVFPPEVSDQELNEINQFLGPEKFEVDSRKIDQSGKIPDELEKLG SLGLGLQVPEYGLGSNTMYSRIGETISMDGSIATVLAHQALGLKGLIAGTEE OKAKYLPKLAGSHIAAFCLTEPAGSDAASIRSATLSDEKKHYLNGSKVWITNGG LANIFVFAKTEVVDGSGVKDKITAFIVERDFGGVTNGPKDGLGRGNTCEVHFE NTKIPVENTILGEVDGDFKVMNINLSGRFSMGVSVAGLLKRLIEMTAETACTRKQFNK

gene

CDS

BASE COUNT 643 a 621 c 692 g 538 t
R L S E F G L I Q B K F A L M A Q K A V M S E M T Y L T A G M L D Q P G F D C S I E A A M K V F S E A A Q Q
C V S E A Q L I G G L Y R D Y P E R L I R O T R I L L I F E G T N E I L M Y A L T G L O H A G R I L T
R I H E L K Q A V T V M D T V G R L R O S L G R T V D L G T G N H V V H P S L A S K A N F E R T Y C F
G R V T E L L R E G K T I M E E Q L V K R V A N I L N G M T A V L S R A S R I G U R N H D H E V L
L A N T F C V E A Y L Q M L F S L S Q D K T A P E N L D E I K K V S Q Q I L E K R A Y T C A H P L D R T C "

[illegible]

RESULT 7
AF327351
LOCUS
DEFINITION

```
ACCESSION AF327351
VERSION AF327351.1 GI:18028282
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2411)
AUTHORS Zhang, J., Zhang, W., Zou, D., Chen, G., Wan, T., Zhang, M. and Cao, X.
TITLE Cloning and functional characterization of ACAD-9, a novel member
of human acyl-CoA dehydrogenase family
JOURNAL Biochem. Biophys. Res. Commun. 297 (4), 1033-1042 (2002)
MEDLINE 2246103
PUBMED 12359260
REFERENCE 2 (bases 1 to 2411)
AUTHORS Zhang, W., Zhang, J., Li, N., Wan, T., Chen, T., Zhang, M. and Cao, X.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2000) Department of Immunology, Second Military
Medical University & Shanghai Brilliance Biotechnology Institute,
800 Xiangyin Rd., Shanghai 200433, P.R. China
LOCATION/Qualifiers
FEATURES
source
1..2411
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
33..1898
/codon_start=1
/product="very-long-chain acyl-CoA dehydrogenase VLCAD"
/db_xref="GI:18028283"
/translation="MSGCGLFLRTTAAARACRLVSVSTANRLRLTSPVRAFAKELF
LGKIKKEVFPPEPVSQDELNEINQFLGVEKFFVEVDNRKIDQEKTPDETLEKLL
SLGLFQVLPSEGGHIAFCITPAGSDAASIRSRATLSDDKKHYILNGSKVITNGG
QKAKYLPKLAEGHIAFCITPAGSDAASIRSRATLSDDKKHYILNGSKVITNGG
LANIFVAKTEVVDSDGSVKKIFAFIVERDFGVNKGPKDLGRSNTCEVHFE
NTKIPENILGEVDFKVMNINLSRGESVYAGLKLRLIEMTAFAEACRKFQFNK
RISEFLIQEKFALMAQKAYVMSMTYLTAGMDQFPDPCSIEAAWVAFVSEAAWQ
CSEVLQILGGIYDYPYRILDRITLLIPETNEILRMVYIALTQLOHAGRILTT
RVHELQAKVSTVMDTVGRRLDRSLGRTVDLGLTNGHVHPSLADSNKFNFTYCF
GRVTEVLLRFNGTMEEOVLKRVANILINLYGMTAVLSRASRSIRIGLRNHDHEVL
LANFVCEAYLQNLFLSLQDLKYAPENLDEQIKKVSQILEKRAYICAHPLDRTC"
BASE COUNT 612 a 604 c 669 g 526 t
ORIGIN
Query Match 99.7%; Score 1858.2; DB 9; Length 2411;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1860; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGAGCGGCTGGGGCTCTTCCTGCGCACACGCGTGGCGCTGCTGCGTCCGGGGCTG 60
Db 33 ATGAGCGGCTGGGGCTCTTCCTGCGCACACGCGTGGCGCTGCTGCGTCCGGGGCTG 92
Qy 61 GTGCTCTACCGCGAACCGCGGCTACTCGGCACACGCGCGCTGTACGAGCTTCGCG 120
Db 93 GTGCTCTACCGCGAACCGCGGCTACTCGGCACACGCGCGCTGTACGAGCTTCGCG 152
Qy 121 AAAGAGCTTTTCTTAGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGC 180
Db 153 AAAGAGCTTTTCTTAGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGC 212
Qy 181 CAAGATGAACCTAATGAATCAATCAGTTCTTGGACCCCGTGGAAAAATTTCTCACTG 240
Db 213 CAAGATGAACCTAATGAATCAATCAGTTCTTGGACCCCGTGGAAAAATTTCTCACTG 272
Qy 241 GAGTGGACTCCCGAAAAATGACAGAGAGGAAAAATCCAGATGAATGAAATTTGGAGAA 300
Db 273 GAGTGGACTCCCGAAAAATGACAGAGAGGAAAAATCCAGATGAATGAAATTTGGAGAA 332
Qy 301 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAAGAGATATGCTGGCGTGGCTC 360
Db 333 TTGAAGAGCTTAGGGCTTTTGGGCTGCAAGTCCCAAGAGATATGCTGGCGTGGCTC 392
Qy 361 TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCCTGTG 420
393 TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCCTGTG 452
421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG 480
453 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG 512
481 CAGAAAGCAAATACTTTGCCCTAAACTGGCGTCCGGGAGACACATTTGAGCCTTCTG 540
513 CAGAAAGCAAATACTTTGCCCTAAACTGGCGTCCGGGAGACACATTTGAGCCTTCTG 572
541 ACGGAGCCAGCAGTGGGAGCGATGCGAGCTCAATCCGGAGCAGAGCAGCAGTAAAGT 600
573 ACGGAGCCAGCAGTGGGAGCGATGCGAGCTCAATCCGGAGCAGAGCAGCAGTAAAGT 632
601 GACAAAGACACTACATCTCAATGGCTCCAAAGTCTGGATTACTTAATGGAGGAGTGG 660
633 GACAAAGACACTACATCTCAATGGCTCCAAAGTCTGGATTACTTAATGGAGGAGTGG 692
661 AATATTTTACTGTGTTTGGCAAAGACTGAGTGGCTGTTGATCTGATGGATCAGTGA 720
693 AATATTTTACTGTGTTTGGCAAAGACTGAGTGGCTGTTGATCTGATGGATCAGTGA 752
721 AATATTTTACTGTGTTTGGCAAAGACTGAGTGGCTGTTGATCTGATGGATCAGTGA 780
753 AATATTTTACTGTGTTTGGCAAAGACTGAGTGGCTGTTGATCTGATGGATCAGTGA 812
781 GATAAATTTAGGCATTCGGGGCTCCAAACACTTTGTAAGTCCATTTTGAAGAACCA 840
813 GATAAATTTAGGCATTCGGGGCTCCAAACACTTTGTAAGTCCATTTTGAAGAACCA 872
841 CCGTGTGAAACATCTTTGGAGAGTCCGAGATGGTAAAGTGGCGCATGAACATCTCT 900
873 CCGTGTGAAACATCTTTGGAGAGTCCGAGATGGTAAAGTGGCGCATGAACATCTCT 932
901 AACAGCGCGGCTTCAGCATGGCGACGCTGCTGGCTGGCTGCTCAAGAGATTGATGA 960
933 AACAGCGCGGCTTCAGCATGGCGACGCTGCTGGCTGGCTGCTCAAGAGATTGATGA 992
961 ATGACTGCTGAGTACGCTGCACAGGAAACAGTTTAAACAGAGCTCAGTGAATTTGA 1020
993 ATGACTGCTGAGTACGCTGCACAGGAAACAGTTTAAACAGAGCTCAGTGAATTTGA 1052
1021 TTGATTTCAGGAGAAATTTTGCATGTGCTCAGAGGCTTACGTATGAGAGATGATGAC 1080
1053 TTGATTTCAGGAGAAATTTTGCATGTGCTCAGAGGCTTACGTATGAGAGATGATGAC 1112
1081 TACCTCACAGCAGGAGTGTGGACCAACCTGGCTTCCCGACTGCTCCATCAGGAGC 1140
1113 TACCTCACAGCAGGAGTGTGGACCAACCTGGCTTCCCGACTGCTCCATCAGGAGC 1172
1141 ATGGTCAAGGTTTACGCTCCGAGCGCGCTGGCAGTGTGAGTGGAGGCTGCGAGATC 1200
1173 ATGGTCAAGGTTTACGCTCCGAGCGCGCTGGCAGTGTGAGTGGAGGCTGCGAGATC 1232
1201 CTGCGGGGCTTGGGCTTACACAGGGGACTATCCGTACGAGCGCATCTGGTGACACCC 1260
1233 CTGCGGGGCTTGGGCTTACACAGGGGACTATCCGTACGAGCGCATCTGGTGACACCC 1292
1261 ATCCTCCTCATCTTCAGGGAACCAATGAGATTCCTCGGATGTACATCCCTGACGGGT 1320
1293 ATCCTCCTCATCTTCAGGGAACCAATGAGATTCCTCGGATGTACATCCCTGACGGGT 1352
1321 CTCAGCATGCCGGCGCATCTCTGACTACCAAGATCCATGAGCTTAAACAGGCAAAAGT 1380
1353 CTCAGCATGCCGGCGCATCTCTGACTACCAAGATCCATGAGCTTAAACAGGCAAAAGT 1412
1381 AGCACAGTCAATGATACCGTTGGCGGAGGCTTCGGAGTCCCTGGCGCAACTGTGGAC 1440
1413 AGCACAGTCAATGATACCGTTGGCGGAGGCTTCGGAGTCCCTGGCGCAACTGTGGAC 1472
1441 CTGGGCTGACAGGCAACCATGAGTTGTCACCCAGTCTTGGGAGAGTGGCAACAAG 1500
```

Db 1473 CTGGGGCTGACAGGCAACACCTGGAGTTGTGCACCCAGTCTTGGGACAGTGCCACAAG 1532
 Qy 1501 TTTGAGGAGAACACTACTGCTTCGGCGCGGACCGTGAGACACTGCTGCTCCGCTTTGGC 1560
 Db 1533 TTTGAGGAGAACACTACTGCTTCGGCGCGGACCGTGAGACACTGCTGCTCCGCTTTGGC 1592
 Qy 1561 AAGACCATCATGAGGAGAGCTGGTACTGAAGCGGTGGCCAAACATCTCTCATCAACCTG 1620
 Db 1593 AAGACCATCATGAGGAGAGCTGGTACTGAAGCGGTGGCCAAACATCTCTCATCAACCTG 1652
 Qy 1621 TATGGCATGACGCGCTGCTGCGCGGCGAGCGCTCCATCCGCATTTGGGCTCCGGAAC 1680
 Db 1653 TATGGCATGACGCGCTGCTGCGCGGCGAGCGCTCCATCCGCATTTGGGCTCCGGAAC 1712
 Qy 1681 CAGCACACAGAGTTCTCTTGGCCAAACACCTTCTGCGTGGAACTTACTTGCAGATCTC 1740
 Db 1713 CAGCACACAGAGTTCTCTTGGCCAAACACCTTCTGCGTGGAACTTACTTGCAGATCTC 1772
 Qy 1741 TTCAGCCTCTCAGCTGGCAAGTATGTCTCAGAAAACCTAGATGAGCAGATTAAGAAA 1800
 Db 1773 TTCAGCCTCTCAGCTGGCAAGTATGTCTCAGAAAACCTAGATGAGCAGATTAAGAAA 1832
 Qy 1801 GTGTCCAGCAGATCTTGGAGAGCGAGCTATATCTGTGCCACCCCTTGGACAGGACA 1860
 Db 1833 GTGTCCAGCAGATCTTGGAGAGCGAGCTATATCTGTGCCACCCCTTGGACAGGACA 1892
 Qy 1861 TGC 1863
 Db 1893 TGC 1895

RESULT 8

BD157796 2400 bp DNA linear PAT 17-JAN-2003
 LOCUS Primer for synthesizing full-length cDNA and use thereof.
 DEFINITION
 ACCESSION BD157796
 VERSION BD157796.1 GI:27863554
 KEYWORDS JP 2002191363-A/12639

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2400)

REFERENCE

1 Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 12639 09-JUL-2002;

TITLE

JUN1

JOURNAL

HELIOS RESEARCH INSTITUTE

COMMENT

OS Homo sapiens (human)
 PN JP 2002191363-A/12639
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KETICHI NAGAI, TETSUJI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof PH Key
 Location/Qualifiers
 FT CDS (498) ..(1991).

FEATURES

SOURCE

1..2400
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT

570 a 629 c 685 g 516 t

ORIGIN

Query Match 89.7%; Score 1671.4; DB 6; Length 2400;
 Best Local Similarity 92.9%; Pred. No. 0;

Matches 1807; Conservative 0; Mismatches 56; Indels 82; Gaps 2;
 Qy 1 ATGAGCGGTGGGGGCTCTTCTTGCACACACCGCTCGGCTCGTCCCTCCGCGGCTCTG 60
 Db 47 ATGAGCGGTGGGGGCTCTTCTTGCACACACCGCTCGGCTCGTCCCTCCGCGGCTCTG 106
 Qy 61 GTGGTCTCTACCGCGAACCGGGGCTTACTGGCCACACCGCGGCTTGTAGAGCTTTTCGCC 120
 Db 107 GTGGTCTCTACCGCGAACCGGGGCTTACTGGCCACACCGCGGCTTGTAGAGCTTTTCGCC 166
 Qy 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAA----- 153
 Db 167 AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAA----- 226
 Qy 154 -----GAAGTTTTCCCATTTTCCAG 172
 Db 227 TCTGGTCCCGCTTTTTCACCCCTCAGCTGCAAGACTGGTGTGAACCTTTGTGAGATTCCCC 286
 Qy 173 AAGTTAGCCCAAGATGA-----ACTTAAATGAATCAATCAGTTCTTGGGACC 218
 Db 287 AAACCTGCCAGAGAGATACACCCCTGGGCGGAGGGGTGTTAAACCTCCGGGATTTCTGAGT 346
 Qy 219 COTGAAAATTTCTTCACTGAAGAGTGGACTCCGAAAAATTTGACAGGAAGGAAAAAT 278
 Db 347 TCCAGAAAACCTTCCAGAGAAAGGTGGACTCCCGAAAAATTTGACAGGAAGGAAAAAT 406
 Qy 279 CCCAGATGAACCTTTGGAGAAATTTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAAG 338
 Db 407 CCCAGATGAACCTTTGGAGAAATTTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCCAAG 466
 Qy 339 AGAATATGGTGGCTGGGCTTCTCCAACACCATGTACTCAAGACTAGGGAGATCATCAG 398
 Db 467 AGAATATGGTGGCTGGGCTTCTCCAACACCATGTACTCAAGACTAGGGAGATCATCAG 526
 Qy 399 CATGATGGGTCCATCACTGTGACCCCTGGCAGCGCACCAGGCTATTGGCTCAAGGGAT 458
 Db 527 CATGATGGGTCCATCACTGTGACCCCTGGCAGCGCACCAGGCTATTGGCTCAAGGGAT 586
 Qy 459 CATCTTGGCTGGCAGTGGAGAGCAAGCAAAATCTTGGCTAAACTGGGCTCCGGGGA 518
 Db 587 CATCTTGGCTGGCAGTGGAGAGCAAGCAAAATCTTGGCTAAACTGGGCTCCGGGGA 646
 Qy 519 GCACATTCAGCCCTTCTGCTCAGCGAGCCAGCCAGTGGGAGCGATGAGCGCTCAATCCG 578
 Db 647 GCACATTCAGCCCTTCTGCTCAGCGAGCCAGCCAGTGGGAGCGATGAGCGCTCAATCCG 706
 Qy 579 GAGCAGCGCCACACTAAGTGAAGCAAGACGACTACATCTCAATGGCTCCAGGCTCTG 638
 Db 707 GAGCAGCGCCACACTAAGTGAAGCAAGACGACTACATCTCAATGGCTCCAGGCTCTG 766
 Qy 639 GATTACTAATGGAGACTGGCCAAATATTTTACTGTGTGTTTCAAGACTGAGGCTCGTTGA 698
 Db 767 GATTACTAATGGAGACTGGCCAAATATTTTACTGTGTGTTTCAAGACTGAGGCTCGTTGA 826
 Qy 699 TTTGATGGATCAGTGAAGACAAATACAGCATTCATAGTAGAAAAGAGACTTTGGTGG 758
 Db 827 TTTGATGGATCAGTGAAGACAAATACAGCATTCATAGTAGAAAAGAGACTTTGGTGG 886
 Qy 759 AGTCACATAATGGGAAACCCGAGATTAATTTAGGCATTTCCGGGCTCCAAACACTTTGAGT 818
 Db 887 AGTCACATAATGGGAAACCCGAGATTAATTTAGGCATTTCCGGGCTCCAAACACTTTGAGT 946
 Qy 819 CCATTTTGAACACCAAGATACCTGTGGAAAACATCTTTGGAGAGCTCGGAGATGGGTT 878
 Db 947 CCATTTTGAACACCAAGATACCTGTGGAAAACATCTTTGGAGAGCTCGGAGATGGGTT 1006
 Qy 879 TAAGGTGGCATGAACATCTTCAACACGCGCGGTTTCCAGCATGGGCGGCTCGTGGCTGG 938
 Db 1007 TAAGGTGGCATGAACATCTTCAACACGCGCGGTTTCCAGCATGGGCGGCTCGTGGCTGG 1066
 Qy 939 GCTGCTCAAGAGATTGATTGAATGACTGCTGAGTAGCCCTGCACAGAGGAACAGTTTAA 998
 Db 1067 GCTGCTCAAGAGATTGATTGAATGACTGCTGAGTAGCCCTGCACAGAGGAACAGTTTAA 1126

```
QY 999 CAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATTTGACATGATGGCTCAGAAGGC 1058
Db 1127 CAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATTTGACATGATGGCTCAGAAGGC 1186
QY 1059 TTACGTCATGAGAGATGATGACCTACCTACAGCAGGATGCTGGAGCAACCTGGCTTTCG 1118
Db 1187 TTACGTCATGAGAGATGATGACCTACCTACAGCAGGATGCTGGAGCAACCTGGCTTTCG 1246
QY 1119 CGACTGCTCCATCGAGGCGCCATGCTGAAGTGTTCACCTCCGAGGCCGCTGGCAGTG 1178
Db 1247 CGACTGCTCCATCGAGGCGCCATGCTGAAGTGTTCACCTCCGAGGCCGCTGGCAGTG 1306
QY 1179 TGTGAGTGAAGGCTCAGATCTCGGGGCTTGGGCTACAAAGGACTATCCGTACGA 1238
Db 1307 TGTGAGTGAAGGCTCAGATCTCGGGGCTTGGGCTACAAAGGACTATCCGTACGA 1366
QY 1239 GCGCATACGCTGACACCCGATCTCTCTATCTTCGAGGGAACCAATGAGATTCGCG 1298
Db 1367 GCGCATACGCTGACACCCGATCTCTCTATCTTCGAGGGAACCAATGAGATTCGCG 1426
QY 1299 GATGTACATCGCCCTGACGGTCTGAGCATGCGCGCGCATCTCACTACAGGATCCA 1358
Db 1427 GATGTACATCGCCCTGACGGTCTGAGCATGCGCGCGCATCTCACTACAGGATCCA 1486
QY 1359 TGAGCTTAACAGAGCCAAAGTAGAGCACAGTCACTGATACCGTTGGCCGAGGCTTCGGGA 1418
Db 1487 TGAGCTTAACAGAGCCAAAGTAGAGCACAGTCACTGATACCGTTGGCCGAGGCTTCGGGA 1546
QY 1419 CTCCTGGGCGGAACCTGTGACCTCGGGCTGACAGCAACCATGAGTGTGCAACCCAG 1478
Db 1547 CTCCTGGGCGGAACCTGTGACCTCGGGCTGACAGCAACCATGAGTGTGCAACCCAG 1606
QY 1479 TCTTGGGACAGTGCCAAAGTTTGAGAGAACACCTACTGCTTCGGCGGACCGTGA 1538
Db 1607 TCTTGGGACAGTGCCAAAGTTTGAGAGAACACCTACTGCTTCGGCGGACCGTGA 1666
QY 1539 GACACTGCTCTCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGT 1598
Db 1667 GACACTGCTCTCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGT 1726
QY 1599 GGGCAACATCTCATCAACTGTATGCGATGACGGCGTCTCTCGGGGCCAGCGCTC 1658
Db 1727 GGGCAACATCTCATCAACTGTATGCGATGACGGCGTCTCTCGGGGCCAGCGCTC 1786
QY 1659 CATCCGATTTGGGCTCGCAACACCGACGAGGTTCTCTTGGCCAAACCTTCTGGGT 1718
Db 1787 CATCCGATTTGGGCTCGCAACACCGACGAGGTTCTCTTGGCCAAACCTTCTGGGT 1846
QY 1719 GGAAGCTTACTTGAGAAATCTTTCAGCTCTCTCAGCTGGCAAGTATGCTCCAGAAA 1778
Db 1847 GGAAGCTTACTTGAGAAATCTTTCAGCTCTCTCAGCTGGCAAGTATGCTCCAGAAA 1906
QY 1779 CTTAGATGAGCAGATTAAAGAGTGTCCAGCAGATCTTTCGAGAGGAGCCTATATCTG 1838
Db 1907 CTTAGATGAGCAGATTAAAGAGTGTCCAGCAGATCTTTCGAGAGGAGCCTATATCTG 1966
QY 1839 TGCCCAACCTCTGTGACAGGACATGC 1863
Db 1967 TGCCCAACCTCTGTGACAGGACATGC 1991
```

```
RESULT 9
AK022568
LOCUS
DEFINITION Homo sapiens cDNA FLJ12506 fis, clone NT2RM2001700, weakly similar
to ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC (EC 1.3.99.-).
ACCESSION AK022568
VERSION AK022568.1 GI:10434032
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE
AUTHORSTITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL

COMMENT

FEATURES
Source

```
1..2400
Location/Qualifiers
1..2400
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="NT2RM2001700"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM2"
/note="cloning vector: pME18SFL3-mRNA from uninduced NT2
neuronal precursor cells."
498..1994
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14104.1"
/db_xref="GI:10434032"
/translation="MYSRLGEIISMDGSTVTTLAAHOAIGLKGILLAGTEOKAKYLP
KLAGSHIAAFCILTEPASGSDAASRATLSLSDKHVILNGSKVMTITNGLIANTFTV
FAKTEVVDSDGSKDKITAFIVERDFGVTNGKPKDKLGIRSNCEVHFNTKIPVE
NLGVEVGDGKVMNINLNSFRSMGVSAGLKLRLIEMTAETACTRKQNKRLSEFGL
IOEKALMAQKAYVMESMTYLAGMDLPDGFPCSEAAAVKVFSEAAQVCSEALQ
ILGGLGYRDPYERILTRILLIPEGTNEILRMVIALTLQHAGRIITTRHLEKQ
AKVSTVMDTVGRRLRSLGTLVDLGTGNHGVHPVSLADSAKFENFTYCFGRVETL
LLRECKTMEEOILVKRVANILINLYGMTAVLSRASRSTRIGLRNHDHVLANTFCV
EAYLQNLFSLSQLDKYAPENLDEQIKVSOQIILEKRAYICAHPLDRTC"
```

CDS

```
BASE COUNT 570 a 629 c 685 g 516 t
ORIGIN
Query Match 89.7%; Score 1671.4; DB 9; Length 2400;
Best Local Similarity 92.9%; Pred. No. 0;
Matches 1807; Conservative 0; Mismatches 56; Indels 82; Gaps 2;
```

```
QY 1 ATGAGCGGCTGCGGGCTCTTCTCGCGCACACCGCTCGGCTCGTCCCTGCCGGGCTCTG 60
Db 47 ATGAGCGGCTGCGGGCTCTTCTCGCGCACACCGCTCGGCTCGTCCCTGCCGGGCTCTG 106
QY 61 GTGGTCTCTACCGCGAAGCGGCGGTACTGCGCACCGCCGCTGTACGAGCTTTCGCC 120
Db 107 GTGGTCTCTACCGCGAAGCGGCGGTACTGCGCACCGCCGCTGTACGAGCTTTCGCC 166
QY 121 AAAGAGCTTTTCTTAGGCAAAATCAAGAGAAA----- 153
Db 167 AAAGAGCTTTTCTTAGGCAAAATCAAGAGAGGTAAACGCGAGCCCTGGCGAACCTTGTCTG 226
QY 154 ----- 172
Db 227 TCTGGCTCCGCTTTTACCCCTCAGCTGCAAGACTGTTGTAACCTTTGTGAGATTTCCAG 286
QY 173 AAGTTAGCCCAAGATGA-----ACTTAATGAATCAATCAGTCTTGGGACC 218
Db 287 AAACCTGCCAGAGAGATACACCCCTGGCGCGGAGCGGTGTTAACACTCCCGATTCTCTGAGT 346
```

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1
Isoqai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y.,
Matsumura, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y. and Sasaki, N.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2400)
Isoqai, T. and Otsuki, T.
Direct Submission
Submitted (23-AUG-2000) Takao Isoqai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'-3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.

QY	219	CGTGGAAAAATCTTTCAC	TGAAGAGGTG	GACTCCCGAAAAATTTG	ACCAGGAAGGAAAT	278
DB						
QY	347	TCCAGGAAAAACCTTCC	CAGAGAAAGTG	GACTCCCGAAAAATTTG	ACCAGGAAGGAAAT	406
DB						
QY	279	CCAGATGAACCTTTGG	GAGAAATG	GAAGAGCCTAGGCTTTT	TGGGCTGCAAGTCCCAGA	338
DB						
QY	407	CCAGATGAACCTTTGG	GAGAAATG	GAAGAGCCTAGGCTTTT	TGGGCTGCAAGTCCCAGA	466
DB						
QY	339	AGAATATGTG	GGCTTCTTCC	AAACCACTG	TACTCAAGACTAGGGAGATCATCAG	398
DB						
QY	467	AGAATATGTG	GGCTTCTTCC	AAACCACTG	TACTCAAGACTAGGGAGATCATCAG	526
DB						
QY	399	CATGATGGTCC	ATCAGCTGT	CACCTG	CACAGCAGCAGAGGTATTTGGGCTCAAGGGAT	458
DB						
QY	527	CATGATGGTCC	ATCAGCTGT	CACCTG	CACAGCAGCAGAGGTATTTGGGCTCAAGGGAT	586
DB						
QY	459	CATCTTGGCT	TG	CGACTGAGGAGCAG	AAAGCCAAATACTTGCCTAAACTTGGGCTCCGGGA	518
DB						
QY	587	CATCTTGGCT	TG	CGACTGAGGAGCAG	AAAGCCAAATACTTGCCTAAACTTGGGCTCCGGGA	646
DB						
QY	519	GCACATTCAG	CGCTTCTGCC	TCACGGAGCAG	CCAGTGGGAGCGATGACGCTCAATCCG	578
DB						
QY	647	GCACATTCAG	CGCTTCTGCC	TCACGGAGCAG	CCAGTGGGAGCGATGACGCTCAATCCG	706
DB						
QY	579	GAGCAGAGCC	ACACTAA	GTGAAGACNAGA	GCACACTACATCTCAATGCTCCAAGTCTG	638
DB						
QY	707	GAGCAGAGCC	ACACTAA	GTGAAGACNAGA	GCACACTACATCTCAATGCTCCAAGTCTG	766
DB						
QY	639	GATTACTAAT	GAGGAC	TGGCCAAATTTTT	TACTGTGTTC	698
DB						
QY	767	GATTACTAAT	GAGGAC	TGGCCAAATTTTT	TACTGTGTTC	826
DB						
QY	699	TTCTGATGG	ATCAGTGA	AAAGCAAAATC	CACAGCATTCATCTAGAAAGAGACTTTGGTGG	758
DB						
QY	827	TTCTGATGG	ATCAGTGA	AAAGCAAAATC	CACAGCATTCATCTAGAAAGAGACTTTGGTGG	886
DB						
QY	759	AGTCACTAAT	TGGGAACCCG	GAAGATAAAT	TAGGCATTGGGGCTCCCAACACTTGTGAAGT	818
DB						
QY	887	AGTCACTAAT	TGGGAACCCG	GAAGATAAAT	TAGGCATTGGGGCTCCCAACACTTGTGAAGT	946
DB						
QY	819	CCATTTTGA	AAACACCA	AGATACCTGT	TGGA	1006
DB						
QY	947	CCATTTTGA	AAACACCA	AGATACCTGT	TGGA	1006
DB						
QY	879	TAAAGTGG	CCATGA	ACATCTC	CAACAGCGCCGTTT	938
DB						
QY	1007	TAAAGTGG	CCATGA	ACATCTC	CAACAGCGCCGTTT	1066
DB						
QY	939	GCTGCTCA	AGAGATTG	ATTTGA	ATATGCTGTG	998
DB						
QY	1067	GCTGCTCA	AGAGATTG	ATTTGA	ATATGCTGTG	1126
DB						
QY	999	CAAGAGG	CTCAGTGA	ATTTGG	ATTTGATTC	1058
DB						
QY	1127	CAAGAGG	CTCAGTGA	ATTTGG	ATTTGATTC	1186
DB						
QY	1059	TTAGCTCAT	GGAGATG	ATACCTAC	CTCAGAGGATGCTGG	1118
DB						
QY	1187	TTAGCTCAT	GGAGATG	ATACCTAC	CTCAGAGGATGCTGG	1246
DB						
QY	1119	CGACTG	CTC	ATCGAGCAG	CCATGTTG	1178
DB						
QY	1247	CGACTG	CTC	ATCGAGCAG	CCATGTTG	1306
DB						
QY	1179	TGTGAGT	GAGCGCTG	CAGATCT	CGGGGCTTTGGG	1238
DB						
QY	1307	TGTGAGT	GAGCGCTG	CAGATCT	CGGGGCTTTGGG	1366
DB						
QY	1239	GCGCAT	ACTG	CGTGAC	ACCGCGCATCTCT	1298
DB						
QY	1367	GCGCAT	ACTG	CGTGAC	ACCGCGCATCTCT	1426
DB						
QY	1299	GATGTACAT	CGCCCTC	AGCGGCTT	CGANGATCGCCGCCGCATCTCT	1358
DB						

[illegible]

Db 1154 GATTTCCCGACTGCTCTATCGAGCGCAGCCATGGTGAAGGTGTTAGCTCTGAGGCTGCCT 1213

QY 1172 GGCAGTGTGTGAGTGGAGCGCTGCAGATCCTCGGGGCTTGGCTACACAAAGGACTATC 1231

Db 1214 GGCAGTGTGTGAGCGAGGCTCTGCAGATCCTTGGGGCTTCAGGCTACATGAAGGACTACC 1273

QY 1232 CTGACGAGCGCATATGCTGCAGACCCCGCATCTCTCATCTTCAGGGAACCAATGAGA 1291

Db 1274 CCTATGAGCGCATGCTGCTGATGCCCGATCTCTCTATCTTTCAGGGAACCAATGAGA 1333

QY 1292 TTCTCCGGATGTACATCGCCCTGACGGGTCTGCAGATGCCCGCGCATCTGAGTACACA 1351

Db 1334 TTCTTCGGTGTTCATGTCCTTGACAGGCTCTGCAGATCTGAGCGCATCTGAGTACACA 1393

QY 1352 GGATCCATGAGCTTAAACAGGCGCAAGTGCAGACAGTCAATGATACCTGTCGCGGAGGC 1411

Db 1394 GGATCAAGAGCTTAAAGTGGCATGTGCACACAGTCAATGAGAGGATGTCGCGAAGC 1453

QY 1412 TTCGGGACTCCTGGCGCAACTGTGGACCTGGGCTGCAGGCAACCAATGAGTGTGC 1471

Db 1454 TTCGGGATTTCTTGGGCGCAACTGTGGACCTGGGCTGCAGGCAACCAATGAGTGTGC 1513

QY 1472 ACCCGAGTCTTTCGGGAGTGCACCAAGTTTTCAGGAGCAACCTACTGCTTCGCGCGGA 1531

Db 1514 ACCCGAGTCTTTCGGGAGTGCACCAAGTTTTCAGGAGCAACCTACTGCTTCGCGCGGA 1573

QY 1532 CGGTGGAGACTGCTGCTCGCTTTGGCAAGACCATGATGAGGAGGAGCTGCTGCTGCTG 1591

Db 1574 CTGTTGAGACCTGTTACTCGCTTTTGGAAAGAACATGTAGAGGAACAGCTGGTGTAA 1633

QY 1592 AGCGGGTGGCAACATCTCATCAACCTGTATGGCATGAGGCGCTGCTGCTGCGGAGCA 1651

Db 1634 AGCGGGTGGCAACATCTCATCAACCTGTATGGCATGAGGCGCTGCTGCTGCGGAGCA 1693

QY 1652 CGCGTCCATCGCATCGGCTTCCGCTCCCAACACGACGAGTCTCTTGGCCACACCT 1711

Db 1694 CGCGTCCATCAGATCGGCTCAGGATCATGATCAGAGGCTTCTGTTGCCACATGT 1753

QY 1712 TCTCGTGGAGCTTACTTGGAGATCTCTTCAGCTCTCTCAGTGGACAGTATGCTC 1771

Db 1754 TCTGTGTGGAGCTTATTTCCAGATCTCTTCAGCTCTCTCAGTGGACAGTATGCTC 1813

QY 1772 CAGAAACCTAGATGAGCAGATTAAGAAGTGTCCACAGCATCTTTCAGAGCGGCTC 1831

Db 1814 CAGAAACCTGATGAGCAGATTAAGAAGTGTCTCGCAGATCTTTCAGAGCGGCTC 1873

QY 1832 ATATCTGTGCCACCTCTGAGCAGCATGC 1863

Db 1874 ATATCTGTGCCACCTCTGAGCAGCATGC 1905

RESULT 11

BC032213

LOCUS

Mus musculus RIKEN cdna 2600017P15 gene, mRNA (cdna clone MGC:38234 IMAGE:5323723), complete cds.

ACCESSION

BC032213

VERSION

BC032213.1

KEYWORDS

MGC.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2148)

REFERENCE

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., and Marra, M.A., Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

23388257

12477932

2 (bases 1 to 2148)

Strausberg, R.

Submitted (06-JUN-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@hgrl.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dieterich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tsurgeon, C., Voigt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 56 Row: 1 Column: 12

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis.

Location/Qualifiers

1. 2148

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:38234 IMAGE:5323723"

/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating

ductal carcinoma. 5 month old virgin mouse."

/clone_lib="NCI_CGAP_Mam6"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 2148

/gene="2600017P15R1K"

/note="synonym: 4732402K02"

/db_xref="LocusID:229211"

/db_xref="MGI:1914272"

51. 1928

/codon_start=1

/product="RIKEN cdna 2600017P15"

/protein_id="AAH32213.1"

/db_xref="GI:21595209"

/db_xref="LocusID:229211"

/translation="MSGCVLLSGATAAAAAAASRVLRREFFARRRPLHTLSQSCSFA

KELFLGNIKQGVFPPEVSHELSEINQVGLPEKFFTEEDSRDQEGKIPVDTL

ELKLSGLGFIQVPEYGGVLSNTMYARLGEITISLSDASITVTTLAAHQAIKLGKILV

gene

CDS

FEATURES

source

GNDEOKAKYLPLKLSGEHIAAFCLTEPAGSDAASIQTRATLSDEKFKYILNGSKWVI TNGGLANFTVFAKTEVVDGSKDKMTAFIVBRDFGCTNGKPEDLIGRSGNCTE VHFNRTPVFNLEVEGVGPKVANNILNSGFNCSAVAGMKLLIELTAYACTRK QFNRLNSFGLIQERFALMAQKAYVNESMAYLTSGMLDQPFDPDCSIEAAMVKVFSSE AAWQCSALQILGSGGYMKDYPIYERMLRDARLILIFEGTNEILRLFTALTGLOHAGR ILTSRIKELSGNVTVMETIGKRLRSLGRTVDLGLDGVHPHSIGDSANKLEEN VHYFRTVETLILREKKNIVBEQLVKRVANILINLYGMTAVLSRSRISIRIGLNHD HEVLLANMFCVEAYEQNLFSLSQLDKNAPENLDEQIKKVSQIILEKRAYICAHPLDRA S*	554 a	504 c	593 g	497 t
BASE COUNT				
ORIGIN				
Query Match	74.1%;	Score 1380.8;	DB 10;	Length 2148;
Best Local Similarity	84.6%;	Pred. No. 0;		
Matches 1550;	Conservative	0;	Mismatches 282;	Indels 0;
Gaps	0;			
QY	32	CGGCTGCGGCTCGTGGCTCGCGGGTCTGGTGGTCTCTACCCGGAACCGCGGCTACTGCG	91	
DB	94	CGGCTGCGGCGGAGCTTCCCGGCTCTGAGGGAATTCACGGCGAGGCGCCCGCGCTGC	153	
QY	92	GCACACGCGCGCTGTACGAGCTTTCCGCCAAAGAGCTTTTCTAGGCAAAATCAAGAAGA	151	
DB	154	ACACAGCCTGCAGAGCTGCTGTTGCCCAAGAGCTTCTCTCGGCAACATCAACGAGA	213	
QY	152	AAGAAGTTTCCCATTTCCAGAAAGTTAGCAAGATGAATTAATGAATCAATAGTTCT	211	
DB	214	AAGGAGTTTCCCATTTCCAGAGTCACTCAACATGAATTAATGAATCAATCAATAGTTCTG	273	
QY	212	TGGGACCCCTGGAAATTTCTACAGAGAGTGGACCTCCGCAAAATTCACGAGGAAG	271	
DB	274	TGGGACCCCTGGAAATTTCTACAGAGAGTGGACCTCTCGAATAATTCACGAGGAAG	333	
QY	272	GGAATATCCAGATGAATTTGGGAAATTTGAAGAGCCTTAGGCTTTTGGGCTGCAAG	331	
DB	334	GAAATATCCAGTTGACACTTAGAGAGTTGAAGAGCCTTGGGACTTTTGGCATACAGG	393	
QY	332	TCCGAGAAATATGTGGCTGGGCTTCTCCACACCATATGACTCAAGACTAGGGGAGA	391	
DB	394	TCCGAGAAATATGTGGCTGGGCTTCTCCACACCATATGACTTGGCTGGGCGGAGA	453	
QY	392	TCATCAGCATGGATGGTGCATCACTGTGACCTTGGCAGCGACACAGGCTATTGGCTCA	451	
DB	454	TCATCAGCCTGGATGCTCTATCACCGTGAACCTGGCAGCACACAGGCTATCGGCTCA	513	
QY	452	AGGGGATCATTTGGCTGGGCTGAGGAGAGAAAGCAATATCTTGCCTAACTGGCGT	511	
DB	514	AGGGGATCATTTGGCTGGGCTGAGGAGAGAAAGCAATATCTTGCCTAACTGGCGT	573	
QY	512	CCGGGAGCATTGTCAGCTTCTGCTCAGCGAGCCAGCGAGTGGGAGGATGTCAGCT	571	
DB	574	CTGGGGAACACATCCGAGGCTTCTGCTTACAGAGCCAGCGAGTGGGAGTATGCTGCAT	633	
QY	572	CAATCCGGAGCAGAGCCACACTAAGTGAAGACAAAGACACTACATCTCAATGGCTCA	631	
DB	634	CCATCCAGACACAGCTAGCTTAAAGTGAAGTGAAGTAAAGTAAAGTAAAGTAAAGT	693	
QY	632	AGTCTGGATTAATATGAGGAGCTGGCCAAATATTTTACTTGTTGTTGAAAGACTGAGG	691	
DB	694	AGTCTGGATTAATATGAGGAGCTGGCCAAATATTTTACTTGTTGTTGAAAGACTGAGG	753	
QY	692	TCGTTGATTTGATGATCAGTCAAGAGCAAAATCAAGCATTCATAGTAGAAGAGACT	751	
DB	754	TGGTTGATTTGATGATGTTGAAACACAGCAAAATGACCGCATTCATAGTAGAAGACT	813	
QY	752	TTGGTGGAGTCACTAATGGGAAACCCGAGATAAATTAGGATTTGGGGCTCCCAACTT	811	
DB	814	TCGGCGGAATCACTAATGGGAAACCCGAGATAAATTAGGATTTGGGGCTCCCAACTT	873	
QY	812	GTGAAGTCCATTTTGAACACACCAAGTACTCTGGAACATCTTGGAGAGCTCGGAG	871	
DB	874	GTGAGTCCATTTTGAATAATACAGAGTGGCTCTGGGAAATGTGCTTGGAGAGTGGAG	933	
QY	872	ATGGGTTTAAAGTGGCCATGAACATCTCTCAACAGCGCGCGGTTTTCAGCATGGGAGCGTGC	931	

RESULT 12

BC031137

LOCUS

DEFINITION

BC031137 3071 bp mRNA linear ROD 10-JUN-2003
Mus musculus RIKEN cdna 2600017p15 gene, mRNA (cdna clone MGC:37218
IMAGE:4971848), complete cds.

DB	934	GGGGCTTTAAAGGTAGCCATGAACATCTCTGAACAGTGGCGGATTCAGCATGGCAGTGGT	993	
QY	932	TGGCTGGGCTGCTCAAGAGATTGTTGAAATGACTGCTGAGTACGCTTCGCAACGAAGAAC	991	
DB	994	TGGCCGGGATGCTGAAGAGCTGATCGAATGACTGCTGAGTATGCTGTACGAGGAAC	1053	
QY	992	AGTTTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAAGAGAAATTTGCATGATGGCTC	1051	
DB	1054	AGTTCAACAGGAATCTCAGTGAATTTGGTGTGATTTCAGGAAAAGTTTGGCTCATGGCTC	1113	
QY	1052	AGAAGGCTTACGTATGAGAGATGACTACTACACAGCAGGAGTCTGGACCAACCTG	1111	
DB	1114	AGAAGGATACGTAATGAGAGATAGGCTTACCTCACCTCAGGATGCTAGACCAACAG	1173	
QY	1112	GCTTTCCGACGCTGCTCCATCGAGGAGCAGCATGGTGAAGGTGTTTCAGCTCCGAGCGGCT	1171	
DB	1174	GATTTCCGACGCTCTATCGAGGAGCAGCATGGTGAAGGTGTTTACCTGAGGCTGCT	1233	
QY	1172	GGCAGTGTGAGTGAGGCGCTGAGATCTCGGGGCTTGGGCTACACAGGAGTATC	1231	
DB	1234	GGCAGTGTGAGGAGGCTCTGAGATCTCTGGGGGCTCAGGCTACATGAAGGATAC	1293	
QY	1232	CGTAGAGCGCATGCTGGGTGACACCGCATCTCTCATCTTCAGGAGAAACCAATGAGA	1291	
DB	1294	CTATGAGCGCATGCTGGGTGATGTCGCGCATTTCTCTATCTTTGAGGAGCAATGAGA	1353	
QY	1292	TTCTCCGGATGATACATCGCCCTGACGGGTCTGCAGATGTCGCGGCGCATCTCTGACTACA	1351	
DB	1354	TTCTCCGGTGTTCATTTGCCCTGACAGGCTGCAGCATCTGAGGAGCATCTGACTCGA	1413	
QY	1352	GGATCATCAGCTTAAACAGGCGCAAGTGCAGACAGTATGATACCGTTTGGCGGAGC	1411	
DB	1414	GGATCAAGAGCTTAAAGTGGCAATGTGACCAAGTCAAGAGAGCATGTTGTCGGAAC	1473	
QY	1412	TTCCGGACCTCTGGGCGCAACTGTGGACTGGGCTGACAGCAACCACTACTCTCTGGCCG	1471	
DB	1474	TTCCGGATTTCTGGGCGCAACTGTGGACTGGGCTTAACAGCGCATCTTGGTGTGCTC	1533	
QY	1472	ACCCAGCTTGGGAGAGTGCACAAAGTTTGGAGGAGAACACCTACTCTCTGGCCGGA	1531	
DB	1534	ACCCAGCTTGGGAGAGTGCACAAAGTTTGGAGGAGAACATGTCCTACTTTGGCCGGA	1593	
QY	1532	CGGTGAGACATGCTGCTCCGCTTTGGCAAGACCATCATGAGGAGCAGCTGGTACTGA	1591	
DB	1594	CTGTGAGACCTTGTTACTCCGCTTTTGGAAAGAACATTTAGAGGAGACAGTGTGTTAA	1653	
QY	1592	AGCGGTGCCCAACATCTCATCAACCTGTATGGCATGAGCGCGTGTCTCGGGGCA	1651	
DB	1654	AGCGGTAGCCAAACATCTCATCAACTTGTATGGCATGACTGCTGTCTCGAGAGCCA	1713	
QY	1652	GGGCTTCATCCGCTTGGGCTCCGCAACAGCAGCAGGCTTCTTTGGCCCAACCT	1711	
DB	1714	GGGCTTCATCAGAAATTTGGGCTCAGGAATCATGATCAGGAGTCTGTTGGCCCAATGT	1773	
QY	1712	TCTGCTGGAAGCTTACTTGCAGAAATCTCTTCAGCCTCTCTCAGCTGGACAAGTATGCTC	1771	
DB	1774	TCTGCTGGAAGCTTATTTCCAGAAATCTCTTCAGCCTGTCTCAGCTGGACAAGAGCTC	1833	
QY	1772	CAGAAACCTTAGATGAGCAGATTAAAGAGTGTCCAGCAGATCTCTTGAAGAGCGGCT	1831	
DB	1834	CAGAAACCTTGAATGAGCAGATTAAAGAGTGTCTCGCAGATCTCTTGAAGAGCGGCT	1893	
QY	1832	ATATCTGTCCCACTCTGGAGAGGACATGC	1863	
DB	1894	ATATCTGTCCCACTCTGGAGAGGACATGC	1925	

ACCESSION

BC031137.1 GI:21410393

VERSION

MGC.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 3071)

REFERENCE

STRAUSBERG, R. D., COLLINS, F. S., WAGNER, L. H., DERGE, J. G.,

KLAUSNER, R. L., ZEEBING, B. B., BUETOW, K. H., SCHAEFER, C. F.,

ALTSCHUL, S. F., JORDAN, H., MOORE, T., MAX, S. I., WANG, J., HSIEH, F.,

HOPKINS, R. F., MARUSINA, K., FARMER, A. A., RUBIN, G. M., HONG, L.,

DIATCHENKO, L., SOARES, M. B., BONALDO, M. F., CASAVANT, T. L.,

STAPLETON, M., SCHNEIDER, A. C., GRIMWOOD, J., TOUCHMAN, J. W., GREEN, E. D.,

SCHEIN, J. E., JONES, S. J., AND MARRA, M. A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2388257

TITLE

12477932

JOURNAL

2 (bases 1 to 3071)

MEDLINE

Strausberg, R.

PUBMED

Direct Submission

REFERENCE

Submitted (03-JUN-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 1IA03, Bethesda, MD 20892-2590,

USA

AUTHORS

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabps@mail.nih.gov

REMARK

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>Contact: (Dickson, Mark) mcdepaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

COMMENT

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 59 Row: b Column: 9

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 27369979.

FEATURES

Location/Qualifiers

1..3071

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:37218 IMAGE:4971848"

/tissue_type="Kidney, normal. 5 month old male mouse."

/clone_lib="NCI-CCAP_Kid14"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..3071

/gene="2600017p15Rik"

/note="synonym: 4732402K02"

/db_xref="LocusID:229211"

/db_xref="MGI:1914272"

46..1923

gene

2600017p15Rik

CDS

1..1923

```

/codon_start=1
/product="2600017p15Rik protein"
/protein_id="AAH31137.1"
/db_xref="GI:21410394"
/db_xref="LocusID:229211"
/translocation="MSGCVLLSRGATAAAAAAASRVLRREFTRRRPLHTSLQCSFA
KELFGNIRKQGVFPFPEVSOHLSSEINOFVGLPKFFTEEDVSRKIDQEGKIPVDTL
EKLSGLFGIQVPEEYGGLSNTMYARLGIISLDASITVTLAAHOAGIKUGLILV
GNDEOKAKYLPKLSGEHIAAFLCLTPASGDAASITQATRSLEDDKYPILNGSKWI
TNGGLANITFPAKTEVDSDGSKDKTAFIVERDFGIGITNGKEDKLGIRGNTCE
VHEPRTVPVENLGEVGFVKVANNILSRFSMGSVAGMLKLIELIATYATKRC
QNRNLSEGLIQEAPALMAKAYVNESMAYLTSGLMDQPPDCSIEAAMVYFSSSE
AAQCVCSEALQILGGSGYMKDPIYERMLRDARILLIFETNEILRLFTALDLOHAGR
ILTSRIKELSGNVTVMETIGRKLRSIDGLTGLDGVHPHSLSGDSANKLEEN
VHYFGRTVETLLRFRGNIEBQLVKRVANILINLYGMTAVLSRASIRITGLRNHD
HBVLLANFCEVAYFQNLFSLSQLOKNAPELDEQIKKVSRLKRAYICAHPLDRA
S"
BASE COUNT      831 a      589 c      789 g      762 t
ORIGIN
Query Match      74.1%; Score 1380.8; DB 10; Length 3071;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1550; Conservative 0; Mismatches 282; Indels 0; Gaps 0;
QY 32 CGGCTGCGGCTCGTCCGCGGCGTCTGTTGGTCTCTACCGCGAACCGGCGGTACTGCG 91
Db 89 CGGCTGCGGCGGCGGCTTCCCGGCGTCTGAGGGAATTCACGCGGAGCGCGCGCGTGC 148
QY 92 GCACGAGCGCGGCTGCTAGCAGCTTTTCGCCAAGAGCTTTTCTAGGCAAAATCAAGAAGA 151
Db 149 ACACGAGCGCTGCAGAGCTGCTGTTCCGCCAAGAGCTCTTCTGGGCAACATCAAGCAGA 208
QY 152 AGAAGTTTTCCTCCATTTCCAGAGTTAGCCAGATGAACATTAATGAATCAATCAGTTCT 211
Db 209 AAGAGTTTTCCTCCATTTCCAGAGTTAGCCAGATGAACATTAATGAATCAATCAGTTCT 268
QY 212 TGGGACCGCTGGAAAAATTTCTACTGAAGAGTTGGACTCCCGGAAAAATTCACGAGGAG 271
Db 269 TGGGACCGCTGGAAAAATTTCTACTGAAGAGTTGGACTCTCGAAAAATTCACGAGGAG 328
QY 272 GAAAAATCCAGATGAAACTTTGGAGAAATTTGAAGAGCTTAGGCGCTATTTGGGCTGCAAG 331
Db 329 GAAAAATCCAGTTGACACCTTAGAGAAATTTGAAGAGCTTAGGCGCTATTTGGGCTGCAAG 388
QY 332 TCCAGAGATATATGTTGGCTGCGCTCTCCACACCATGTTACTCAAGACTAGGAGGAGA 391
Db 389 TCCAGAGATATATGTTGGCTGCGCTCTCCACACCATGTTACTCAAGACTAGGAGGAGA 448
QY 392 TCATCAGCATGGATGGGTGCCATCATCTGTGACCCCTGGCAGCGCACAGGCTATTGGCCTCA 451
Db 449 TCATCAGCATGGATGGGTGCCATCATCTGTGACCCCTGGCAGCGCACAGGCTATTGGCCTCA 508
QY 452 AGGGATCATCTTGGCTGCGCTAGGAGCAAGCAATAGTTCCCTAAATGCGCTGCGGT 511
Db 509 AGGGATCATCTTGGCTGCGCTAGGAGCAAGCAATAGTTCCCTAAATGCGCTGCGGT 568
QY 512 CCGGGGAGCATTCAGAGCTTTCTGCTCAGCGGAGCAGCAGTGGGAGGAGTGCAGCCT 571
Db 569 CTGGGGAACACATCGCAGCTTCTGCTTACAGAGCCGCGCAGGAGTGGAGTGTGCTGCAT 628
QY 572 CAATCCGAGCAGAGCCACACTAGTGAAGCAAGAACAGCACTACATCTCAATGGGTCCA 631
Db 629 CCATCCGAGCAGAGCTTACGTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 688
QY 632 AGGTCTGATTAATGAGGAGTGCCTCAATATTTTACTGTGTGTTGCAAGACTGAGG 691
Db 689 AGGTCTGATTAATGAGGAGTGCCTCAATATTTTACTGTGTGTTGCAAGACTGAGG 748
QY 692 TCGTTGATTAATGAGGAGTGCCTCAATATTTTACTGTGTGTTGCAAGACTGAGG 751
Db 749 TCGTTGATTAATGAGGAGTGCCTCAATATTTTACTGTGTGTTGCAAGACTGAGG 808
QY 752 TTGTTGAGTCACTAATGGGAAACCCGAAAGATAAATTAGGCATTCGGGGCTCCAACTT 811

```

Db	809	TCGGCGAATCACTAATGGAAACCGGAGATAAATAGGCATTCAGGCTCCACACAT	868
QY	812	GTGAAGTCCATTTTGAACACACCAAGATACCTGTGGAAACATCTCTGGAGAGTTCGGAG	871
Db	869	GTGAGTCCATTTTGAACATACCAAGTGCCTGTGGAAATGTCTTGGAGAAGTTGGAG	928
QY	872	ATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGGTTCAGCATGGCGCGCTCG	931
Db	929	GGGGCTTTAAGTAGGCATGAACATCCTCAACAGTGGGCGATTCAGCATGGCGCGCTCG	988
QY	932	TGCTGGGCTGCTCAAGAGATTGAATGAATGACTGCTGAGTACGCTGCACAAAGAAAC	991
Db	989	TGCGCGGATGCTGAAGAGCTGATCGAATGCTGCTGAGTATGCTGTACGAGGAAC	1048
QY	992	AGTTTAAAGAGGCTCACTGAATTTGGATTGATTCAGGAGAAATTCAGCATGATGGCTC	1051
Db	1049	AGTTCAACAGGATCTCAGTGAATTTGGTTGATTCAGGAAAGTTTGGCTCATGGCTC	1108
QY	1052	AGAGGCTTACGTCATGGAGATGATGACCTACCTACAGCAGGATGCTGGACCAACCTG	1111
Db	1109	AGAAGGCATACGTAATGGAGATGATGCTTACCTACCTCAGGATGCTAGACCAACAG	1168
QY	1112	GCTTTCCGACTGCTCCATCGAGGACCGCATGTTGAGTGTTCAGTCCGAGGCGCGCT	1171
Db	1169	GATTTCCGACTGCTCTATCGAGGACCGCATGTTGAGTGTTCAGTCCGAGGCGCT	1228
QY	1172	GGCAGTGTGAGTGGGCTGCAGATCTCGGGGCTTGGCTACACAAAGGACTATC	1231
Db	1229	GGCAGTGTGAGGAGGCTCTGCAGATCTTGGGGCTCAGGCTACATGAGGACTACC	1288
QY	1232	CGTACGAGCCATACCTCGCTGACACCCGCTCCTCTCATCTTCGAGGAAACCAATGAG	1291
Db	1289	CTATGAGCCATGCTCGGTGATGCGGCATCTCTCTTAATCTTTGAGGAAACCAATGAG	1348
QY	1292	TTCTCCGATGATACATGCGCTGAGCGGTGTCAGCATGCGCGCCCATCTTCAGTACCA	1351
Db	1349	TTCTCGGTTGTTCAATTCGCTGACAGGCTGACGATGCTGACGATCTTCAGCTCGA	1408
QY	1352	GGATCCATGAGCTTAAACAGGCAAGTGAACAGTATGATGATGATGCTTGGCGGAGGC	1411
Db	1409	GGATCAAGAGCTTAAAGTGGCATGTGACACAGTATGATGATGATGATGCTTGGCGGAA	1468
QY	1412	TTGGGACTCCCTGGGCGCACTGTGAGCTGCGGCTGACAGCAGCAGTGGATGTGTC	1471
Db	1469	TTGGGATCTCTGGGCGCACTGTGAGCTGCGGCTTAAACAGGATCTTGTGTGTGTC	1528
QY	1472	ACCCAGTCTTGGGAGCTGCAACAGTGTGAGGAGAACCTACTGCTTGGCGGGA	1531
Db	1529	ACCCAGCTTGGGAGCTGCAACAGTGTGAGGAGAACCTACTGCTTGGCGGGA	1588
QY	1532	CCGTGGAGACACTGCTGCTCGCTTTGGCAAGACCAATCATGAGGAGCAGCTGCTGTA	1591
Db	1589	CTGTTGAGACCTGTTACTCGTTTGGAAAGAACATGTTAGAGAACAGCTGTGTGTA	1648
QY	1592	AGCGGTGGCCAGACCTCATCAACCTGTATGCGATGACGCGCTGCTCGCGGCA	1651
Db	1649	AGCGGTGAGCAACCTCATCAACCTGTATGCGATGACGCGCTGCTCGCGGCA	1708
QY	1652	GCGCTCATCGCATTTGGGCTGCGCAACAGCAGCAGGATCTCTTGGCGCAACCT	1711
Db	1709	GCGCTCATCGCATTTGGGCTGCGCAACAGCAGGATCTCTTGGCGCAACCT	1768
QY	1712	TCTCGTGGAGAGCTTACTTGCAGAACTCTTTCAGCTCTCTCAGCTGCAGCAAGTATCTC	1771
Db	1769	TCTGTGTGGAAGCTTATTTCCAGAACTCTTTCAGCTCTCTCAGCTGCAGCAAGTATCTC	1828
QY	1772	CAGAAACCTAGATCAGCAGATTAAGAAAGTGTCCAGCAGATCTTTCAGCAAGGAGCT	1831
Db	1829	CAGAAACCTGATGAGCAGATTAAGAAAGTGTCTCGCAGATCTTTCAGCAAGGAGCT	1888
QY	1832	ATATCTGCGCCACCTCTTGACAGGATGC	1863
Db	1889	ATATCTGTGCCCCACCGCTGGACAGCATCC	1920
RESULT 13			
LOCUS	BC001817		
DEFINITION	Homo sapiens, Similar to NP0002 protein, clone MGC:3368		
IMAGE:3530094, mRNA, complete cds.			
ACCESSION	BC001817		
VERSION	BC001817.1	GI:12804754	
KEYWORDS	MGC, Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1494)		
AUTHORS	Straussberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-JAN-2001) National Institutes of Health, Mammalian		
REMARK	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
COMMENT	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
	NIH-MGC Project URL: http://mgc.nci.nih.gov		
	Contact: MGC help desk		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Institute for Systems Biology		
	http://www.systemsbio.org		
	Contact: amadan@systemsbiology.org		
	Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia		
	Greene, Mark Kettelman and Anuradha Madan		
FEATURES	Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
source	Series: IRAL Plate: 11 Row: k Column: 19.		
	Location/Qualifiers		
	1..1494		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="MGC:3368 IMAGE:3530094"		
	/tissue_type="Muscle, rhabdomyosarcoma"		
	/clone_lib="NIH MGC.17"		
	/lab_host="DH10B-R"		
	/note="Vector: pOTB7"		
	475..1041		
	/codon_start=1		
	/product="Similar to NP0002 protein"		
	/protein_id="AAH01817.1"		
	/db_xref="GI:12804755"		
	/translation="MYIALTGLQHAGRIITRIHELKQAKVSTVMDTVGRRRLDSLGR		
	TVDLGTGNHGVVHPVSLADSKANKFENTYCFGTVEITLLRFGKTIMEQLVKRVAN		
	ILINLYGMTAVLSRSTRIRLNHDHVEILLANTFCVEAYLQNLFSLSQDKYAPEN		
	LDEQIKKVSQQLKRAYTCAHPLDRT"		
	BASE COUNT 370 a 386 c 422 g 316 t		
ORIGIN			
	Query Match 55.2%; Score 1029; DB 9; Length 1494;		
	Best Local Similarity 100.0%; Pred. No. 6.7e-238; Indels 0; Gaps 0;		
	Matches 1029; Conservative 0; Mismatches 0		
QY	835	AAGATACCTCTGGAACATCCTTGGAGAGTCGGAGATGGGTTTAAAGTGGCCATGAAC	894
Db	10	AAGATACCTCTGGAACATCCTTGGAGAGTCGGAGATGGGTTTAAAGTGGCCATGAAC	69
QY	895	ATCCTCAACAGCGCGGTTTACAGATGGGCGAGCTGCTGGCTGGCTGCTCAGAGATTG	954
Db	70	ATCCTCAACAGCGCGGTTTACAGATGGGCGAGCTGCTGGCTGGCTGCTCAGAGATTG	129
QY	955	ATTCAATGACTGCTGAGTACGCTGCACAGGAACAGTTTAAACAGAGGCTCAGTGAA	1014

Db 587 ACGGAGCCAGCCAGTGGGAGCGATGCGAGCTCAATCCGAGGAGAGAGCCACACTAAGTGAA 646
 QY 601 GACAAGAGCAGTACATCTCAATGCTCCAGGCTGATTAATGGAGAGTGGCC 660
 Db 647 GACAAGAGCAGTACATCTCAATGCTCCAGGCTGATTAATGGAGAGTGGNC 706
 QY 661 AATATTCTACTGTGTTGCAAGAGTGGGCTGTTGATTTCTGATGATCAGTGAAGAC 720
 Db 707 AATATTCTACTGTGTTGCAAGAGTGGGCTGATTCG-TGATTTCTGATGATCAGTGAAGAC 765
 QY 721 AAAATC 726
 Db 766 AAAATC 771

RESULT 15

AK027186 1995 bp mRNA linear PRI 29-SEP-2000
 LOCUS Homo sapiens cDNA: FLJ23533 fis, clone LNG06931, highly similar to
 DEFINITION AF078854 Homo sapiens NP0002 mRNA.

AK027186 1 GI:10440254
 VERSION oligo capping; fis (full insert sequence).

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Kawakami, T., Nouchi, S., Itoh, T., Shigeta, K., Senba, T.,
 Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Tanigami, A.,
 Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M.,
 Ohmori, Y., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T.,
 Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 NEDO human cDNA sequencing project

TITLE

JOURNAL Unpublished

REFERENCE

AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
 Shibahara, T., Tanaka, T. and Nakamura, Y.
 Direct Submission
 Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
 Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
 Fax: 81-3-5449-5416)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing; Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

FEATURES

source

1..1995

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="LNG06931"

/tissue_type="human lung"

/clone_lib="LNG"

/note="cloning vector pME18SFL3"

1..1995

/note="highly similar to AF078854 Homo sapiens NP0002

misc_feature

mRNA

BASE COUNT 492 a 513 c 553 g 437 t

ORIGIN

Query Match 28.2%; Score 526; DB 9; Length 1995;
 Best Local Similarity 98.2%; Pred. No. 3.7e-116;
 Matches 532; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1031 AGAATTTGCTGATGCTCAGAGGCTTACGTCATGAGAGTATGACCTACCTCAG 1090
 Db 55 AAAATTTGCTGATGCTCAGAGGCTTACGTCATGAGAGTATGACCTACCTCAG 114

QY 1091 CAGGATGCTGACCAACCTTGCTTCCGACTGCTCCATCGAGGAGAGCCATGTTGAAGG 1150
 Db 115 CAGGATGCTGACCAACCTTGCTTCCGACTGCTCCATCGAGGAGAGCCATGTTGAAGG 174
 QY 1151 TGTTGAGTCCGAGCCGCTGAGTGTGAGTGAGGCGCTGCAGATCCTCGGGGGCT 1210
 Db 175 TGTTGAGTCCGAGCCGCTGAGTGTGAGTGAGGCGCTGCAGATCCTCGGGGGCT 234
 QY 1211 TGGGCTACACAAGGACTATCCGTACGAGCGCATACTGGTGACACCCGCAATCCTCCTCA 1270
 Db 235 TGGGCTACACAAGGACTATCCGTACGAGCGCATACTGGTGACACCCGCAATCCTCCTCA 294
 QY 1271 TCTTCAGGAACCAATGAGATTCCTCCGATGTACATGCCCTGACGGGTCTGCAGCATG 1330
 Db 295 TCTTCAGGAACCAATGAGATTCCTCCGATGTACATGCCCTGACGGGTCTGCAGCATG 354
 QY 1331 CCGGCGCATCTGACTACCAAGGATCCATGAGCTTAAACAGGCCAAAGTGAGCAGTCA 1390
 Db 355 CCGGCGCATCTGACTACCAAGGATCCATGAGCTTAAACAGGCCAAAGTGAGCAGTCA 414
 QY 1391 TGGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGACCTGGGGCTGA 1450
 Db 415 TGGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGACCTGGGGCTGA 474
 QY 1451 CAGGCAACCATGGAGTTGTGACCCCGAGTCTTGGGACAGTGCACAAAGTTTGAGGAGA 1510
 Db 475 CAGGCAACCATGGAGTTGTGACCCCGAGTCTTGGGACAGTGCACAAAGTTTGAGGAGA 534
 QY 1511 ACACCTACTGCTTCGGCCGAGCCGTTGGGAGACACTGCTGCTCCGCTTTGGCAAGACCATCA 1570
 Db 535 ACACCTACTGCTTCGGCCGAGCCGTTGGGAGACACTGCTGCTCCGCTTTGGCAAGACCATCA 594
 QY 1571 TG 1572
 Db 595 GG 596

Search completed: September 6, 2003, 17:23:08

Job time : 6856.73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 10:14:13 : Search time 107.074 Seconds
(without alignments)
7679.711 Million cell updates/sec

Title: US-09-945-326-3

Perfect score: 1863

Sequence: 1 atgagcgctcgggctctt.....accctctggacagacatgc 1863

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	168.6	9.0	1341	4	US-09-252-991A-11446
2	165.6	8.9	1212	4	US-09-252-991A-11389
3	155.6	8.4	1173	4	US-09-252-991A-9487
c 4	155.6	8.4	1209	4	US-09-252-991A-9346
c 5	143.4	7.7	4403765	3	US-09-103-840A-2
c 6	143.4	7.7	4411529	3	US-09-103-840A-1
c 7	139.4	7.5	3515	4	US-09-221-017B-824
8	134	7.2	1395	4	US-09-252-991A-16537
c 9	134	7.2	1665	4	US-09-252-991A-16061
10	132.6	7.1	33312	4	US-08-311-731A-121
11	132.2	6.6	1260	4	US-09-252-991A-14526
c 12	123.2	6.6	1260	4	US-09-252-991A-14838
13	111.2	6.0	1155	4	US-09-648-004-5
14	111.2	6.0	17417	4	US-09-648-004-27
15	105.8	5.7	1170	4	US-09-328-352-2254
16	102	5.5	1275	4	US-09-252-991A-8082
17	102	5.5	1488	4	US-09-252-991A-8036
c 18	102	5.5	2055	4	US-09-252-991A-8231
19	97	5.2	1185	4	US-09-328-352-2675
20	92.2	4.9	4403765	3	US-09-103-840A-2
21	92.2	4.9	4411529	3	US-09-103-840A-1
22	91.8	4.9	849	4	US-09-252-991A-12573
c 23	91.8	4.9	969	4	US-09-252-991A-12911
24	91.8	4.9	1257	4	US-09-252-991A-12881
25	90.4	4.7	1539	4	US-09-364-230-29
26	87.2	4.7	1164	4	US-09-328-352-2316
c 27	86.6	4.6	1440	4	US-09-252-991A-2675

28	86.6	4.6	2592	4	US-09-252-991A-3003	Sequence 3003, Ap
29	85.4	4.6	1553	4	US-09-364-230-31	Sequence 31, Appl
30	74.4	4.0	2079	4	US-09-252-991A-11426	Sequence 11426, A
31	74	4.0	1248	4	US-09-328-352-573	Sequence 573, App
32	72	3.9	1209	4	US-09-328-352-155	Sequence 155, App
33	70.8	3.8	1056	4	US-09-364-230-27	Sequence 27, Appl
34	70.2	3.8	1230	4	US-09-252-991A-14780	Sequence 14780, A
35	66.2	3.6	1182	4	US-09-252-991A-6534	Sequence 6534, Ap
c 36	66.2	3.6	1419	4	US-09-252-991A-6499	Sequence 6499, Ap
c 37	60	3.2	513	4	US-09-252-991A-9318	Sequence 9318, Ap
38	57.6	3.1	1149	4	US-09-328-352-724	Sequence 724, App
39	54.8	2.9	276	4	US-09-252-991A-9426	Sequence 9426, Ap
40	53.8	2.9	958	4	US-09-155-183-10	Sequence 10, Appl
41	53.8	2.9	1227	4	US-09-328-352-1449	Sequence 1449, Ap
42	53.4	2.9	766	4	US-09-364-230-33	Sequence 33, Appl
43	53.4	2.9	1200	4	US-09-328-352-234	Sequence 234, App
c 44	52.4	2.8	7218	1	US-08-232-463-14	Sequence 14, Appl
45	50.8	2.7	1302	4	US-09-252-991A-14660	Sequence 14660, A

ALIGNMENTS

RESULT 1

US-09-252-991A-11446/c

; Sequence 11446, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252.991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-07-788

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 11446

; LENGTH: 1341

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; US-09-252-991A-11446

Query Match	9.08;	Score 168.6;	DB 4;	Length 1341;
Best Local Similarity	49.9%;	Pred. No. 5.2e-37;		
Matches 553;	Conservative 0;	Mismatches 529;	Indels 27;	Gaps 4;
Qy	242	AGGTGGACTCCCGAATAATTTGACCAAGGAGGAAATCCCGAGATGAACATTTGAGAAAT	301	
Db	1220	AGCGCGCGCGCGCAGATCGACGAGCAATCAGTTCCCATGGACATGTGGCGCAGT	1161	
Qy	302	TGAAGAGCCTTAGGCGCTTTTGGCTGCAAGTCCCGAAGAAATATGTTGGCTGGCTTCT	361	
Db	1160	TCGGCAGATGGCGCTGCTCGGCATCACCGTCGACGAGGAGTACGCGGTTTCGGCGCTGG	1101	
Qy	362	CCAAACCATGTACTCAAGACTAGGGAGATCATCAGCATGG---ATGGTTCATCACTG	418	
Db	1100	GCTACCTGGCCCGCGCTGGTTCATGAAGAGATCAAGCGGCGCTCGGCTCGGTGGCGC	1041	
Qy	419	TGACCTTGGCAGCGCACAGGCTATTGGCTCAAGGGATCATCTTTGGCTGGCAGCTAGG	478	
Db	1040	TCCTCTATGGCGCGCATTCACACTCTGGTCAACAGATCAAGCGCAACGCGCGC	981	
Qy	479	ACGAAAGCAAAATACTTGCCTAAACTGGCGTCCGGGAGCAGCATTTGACGCTTCTGCC	538	
Db	980	AACAAAGGCGCGCTACCTGCCGCGCTGTATCCGCGCAACACATCGGTGCGCTGCCGA	921	
Qy	539	TCAGGAGCGCAGCAGTGGGAGATGACGCTCAATCCGGAGCAGGACACACTAAGTG	598	
Db	920	TGAGCGAACCACACGCGCGCTCGGACGTGCTCGATGAAGCTGCGCGCGGACCGGTGCG	861	

	Qy	Db
1259	GCATCCTCCTCATCTTCGAGGGACCAATGAGATTCTCCGGATGTACATCGGCC	1312
1133	AGCTCTACGAGATCGGCGCGGCACACGAGATCGCGCGCATGCTGATCGGCC	1186

RESULT 3

```

US-09-252-991A-9487
; Sequence 9487; Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9487
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-9487

```

Query Match	8.4%	Score 155.6	DB 4	Length 1173
Best Local Similarity	50.5%	Pred. No. 2e-33		
Matches 526	Conservative 0	Mismatches 489	Indels 27	Gaps 5
QY	277	ATCCACGATGAACCTTTGGAGAAATTTGAAGACGCTAGGCGCTTTTGGCGCTGCAAGTCCCA	336	
Db	124	ATCCCCGAGGGATCATCGCCCCCATATCGCGGAGATGGGCCCTGTTGGCGCTTCGTATCCG	183	
QY	337	GAGAATATGGTGGCCCTGGGCTTCTCCACACCATGTACTCAAGACTAG--GGGAGATC	393	
Db	184	GAGGCTTACGCGCGGCTCGGGGTGACCATGGAGGAAGAGGTGAGCATCGCCTTCGACGCTG	243	
QY	394	ATCAGCATGATGGGTGCCATCACTGTGACCCCTGGCAGCGCACCGAGCTATTTGGCGCTCAAG	453	
Db	244	GGCGGACCTCGCGCGGCTTCGGTTCGTGCTGGCAGCAACACAGGCATCGGCTCCAG	303	
QY	454	GGGATCATTTGGCTGGCCACTGAGGAGCAGAAAGCAAAATACTTGCCTAAACTGGCGTCC	513	
Db	304	GGCATGTCATCGACGGCACCGAAGAACAGAGCGGCGCTACCTGCGCGCTGGCGCTCG	363	
QY	514	GGGAGCAGCATTCACGCGCTTCGCTCAGCGAGCCAGCCAGTGGAGCGATGACGCTCA	573	
Db	364	GGCGAACTGCTCAGTTGTTGTCGCTCAGCAGAACCGGACTCCGGTTCGCGAGCGCGCTCG	423	
QY	574	ATCCGAGCAGAGCCACACTAAGTGAAGACAGAAAGCACTACATCTCTAATGGCTCCAAAG	633	
Db	424	CTGAAGACCAAGCGCGTGC-----GTGACGCGGAGCACTACGTCTCAAGGTACCAAG	477	
QY	634	GTCTGGATTAATAGGAGGACTGGCCAATATTTTACTGTGTTTGCAGAGACTGAGGTC	693	
Db	478	CGCTTCATCACAAGCCCCGACGGCGGCACTACACCGTGTGTCGCGCGCACCGACCCG	537	
QY	694	GTTGATTCTCGTGGATCAGTGAAGACAAAAATCACAGCATTCATAGTAGAAGAGACTTT	753	
Db	538	GCATCCGTGGCGCC-----GGCGGATCTCCGTTTCGTGCTGAGCGCGGTACG	588	
QY	754	GGTGGAGTCACTAATGGGAACCCGAGATAAATTAGGCATTCGGGCGCTCCAAACATTGT	813	
Db	589	CCGGGCGCTGTCCTTGGGAAGCCGGATCGCAAGATGGGACAGAAGGCGCGCATCCTGT	648	
QY	814	GAGTCCATTTTGAAGAACACCAAGATACCTGTGGAAGAACATCCTTGGAGAGGTC---GGA	870	
Db	649	GACGTGATCTTCGACGACTCGCGGTCCTGCCAGTCAGTTGATCGCGGGGTGAGGCG	708	
QY	871	GATGGGTTAAGGTGGCCATGAACATCCTCAACAGCGCGCGGTTTCAGCATGGGACGCTC	930	

RESULT 4

```

RESOLUTION
US-09-252-991A-9346/c
; Sequence 9346, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9346
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9346

```

Query Match	8.4%	Score 155.6;	DB 4;	Length 1209;
Best Local Similarity	50.5%;	Pred. No. 2.1e-33;		
Matches 526;	Conservative 0;	Mismatches 489;	Indels 27;	Gaps 5;
QY	277	ATCCACAGATGAACACTTTTGAGAAATTTGAAGAGCCTAGGCGCTTTTGGCTGCGAAGTCCCA	336	
Db				
QY	1068	ATCCCGAGCGCATCATCGCCGCATCCGGAGATGGCGCTGTTTCGGCTTTCGATCCCG	1009	
QY	337	GAAGAAATATGTTGGCTCGGGCTTCTCCACACCATCTACTCAAGACTAG---	393	
Db				
QY	1008	GAGGCCATACGGCGGCTCGGGGTGACCATGGAGAAAGGTGAGCATCGCCTTCGAGCTG	949	
QY	394	ATCAGCATGGATGGGTCCCATCTGTGACCTTGGCAGCGCACCGCTTATGGCCTCAAG	453	
Db				
QY	948	GGCGGACCTCGCGGCGTTCGTTGCTGGGCACCAACACGGCATCGGCTCCCG	899	
QY	454	GGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTGGCTAAATGGCGTCC	513	
Db				
QY	888	GGCATCGTCATCGACGGCACCGAAGAACAGACGGCGCTACTGTCGGCGCTTGGCGTGG	829	
Db				

RESULT 5

Query Match	7.7%	Score 143.4	DB 3	Length 4403765
Best Local Similarity	48.2%	Pred. No. 3.2e-28		
Matches 513; Conservative	0	Mismatches 531	Indels 21	Gaps 3
QY 260	TTGACCAAGGAAGGAAAAATCCAGATGAACACTTTGAGAGAAATTTGAAGACCTTAGGCGCTTT	319		
DB 3652814	TGACAGAGAAGCGCGGTTCCCGGAGAAGACGGCTAGTGGCGCTCAATTCCTCCGGTTTCA	3652755		
QY 320	TTGGGCTGCAAGTCCCGAGAAGAAATATGGTGGCTCGGGCTTCTCCACACCATGTACTCAA	379		
DB 3652754	ACGCGTCCACATTCGCGAGAGTACGGCGGTACGGGTCCGACTCGGTAGCTACCTGCA	3652695		
QY 380	GACTAGGGAGATCATCAGCATGATGGGTGCCATCACTGTGACCCCTGGCAGGCGCACGAG	439		
DB 3652694	TCGTGATCGAAGAGTGGCCCGGTGCACCGCTCCGCTGTTGATCCCGCGGTCAACA	3652635		
QY 440	CTATTGGCCCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCAATATCTTGC	499		
DB 3652634	AGCTGGCACCATTGGGCTGATCCTCGGGGCTCCGAGAGCTGGAAGAAGCAGGTGCTGC	3652575		
QY 500	CTAACTGGCGTCCGGGGACACATTGCAACCTTCTGCCTCACGGAGCCAGCGAGTGGGA	559		
DB 3652574	CGGCGTTGGCGCGGAAGGGCGATGGCTTCCATGATTTGATGTAACGGAGCTACATATCC	3652515		
QY 560	GCAGTCAGCCTCAATCCGGAGCAGACACACTAAGTGAAGACAAAGACACTATATCC	619		
DB 3652514	GTGACGGCGGTCCATCGACGACGGGCC-----AAGCCGATGGGATCACTGGATTC	3652461		
QY 620	TCAATGGCTCCAAAGTCTGGATTACTAATGGAGACTGGCCAAATATTTTACTGTGTTTG	679		
DB 3652460	TCAACGGCGCAAGTGTGTGATACCAACGCGGCAAGTCGACGTGACACAGGTGATGG	3652401		
QY 680	CAAAAGACTAGGTCGTTGATTCTGATGATCAGTGAAGACAAATCACAGCAATTCATAG	739		
DB 3652400	CGGTGACCGA-----TCCCGACCGGGCGCCACAGGCATCTCGCGTTTATGG	3652353		
QY 740	TAGAAAGAGACTTTGTTGGAGTCACTAATGGGAACCCGGAATATAATTAGGCATTCGGG	799		
DB 3652352	TGCACAGGACACAGGGGTTACCGTCGTCGGAAGAAGCAAGCTCGGGATCAAGG	3652293		
QY 800	GCTCCACACTTTGTAAGTCCATTTTGAACACCAAGATACCTGTGGAACACATCCCTTG	859		
DB 3652292	GGTCACCGACACCGAGTGTACTTCAGAACTCCCGCATCCCGCGGATGCGCATATCG	3652233		
QY 860	GAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGGTTACGA	919		
DB 3652232	GTGAGCCCGGTACCGGTTTCAAGACCGCGCTGGCCACGTTGGACCAACCCGTCACGA	3652173		
QY 920	TGGCGAGCTGCTGGCTGGCTGTCAAGAGATTGATTGAATGACTGCTGAGTACGCTT	979		
DB 3652172	TTGGGCGCCAGCCGTGGGTATCCCGACGGCGCGCTGGATGCTGCCATCGCCTACACCA	3652113		
QY 980	GCACAAGGAACAGPTTTAACAGAGGCTCAGTGAATTTGGATTGATTTCAGGAGAAATTG	1039		
DB 3652112	AGGACCGCAAGCAGTTCGGTGGTGCATGATGTTTCAGGCGGCTGCAGTCACTGCTGG	3652053		
QY 1040	CACTGATGGCTCAGAAGCTTTACGTCATGGAGAGTATGACCTACTCACACGACGGATGC	1099		
DB 3652052	CCGACATGGCATGAAGTGGAGGGCGTTCGGCTGATGGTTTACTCCGCCCGCC-----C	3651999		
QY 1100	TGGACCAACCTTGGCTTTCCCGACTGCTCCATCGAGGCAAGCCATGGTGAAGTGTTCAGT	1159		

Db 3651995 GCACCGAAGCCGGTGAAGCCGATCTGGGCTTCATTTCCGGCGCTCGAAGTCTTCGCT 3651936
QY 1160 CCGAGGCGGCTGGCAGTGTGTGAGTGGCGCTGCAGATCCCTCGGGGCTTGGGCTACA 1219
Db 3651935 CCACGTCGGATGGAGGTACACCGATGCCGTACAACCTGTTCCGGCGCGCGGCTACA 3651876
QY 1220 CAAGGGCATATCCGTACGAGCGCATACTCGGTGACACCCGCGCTCTCTCATCTTCGAGG 1279
Db 3651875 CCACCGACTTCCCGTTCGAGCGGTTTATCGCGACGCCAAGATCACCCAGATCTATGAGG 3651816
QY 1280 GAACCAATGAGATCTCCCGATGTACATGCCCTGAGCGGTGTC 1324
Db 3651815 GCACCAATGAGATTCAGCGGGTAGTGTATCGCGCGCGCTACTGC 3651771

RESULT 6

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: 24366-20007.00
; CURRENT FILING DATE: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 7.7%; Score 143.4; DB 3; Length 4411529;
Best Local Similarity 48.2%; Pred. No. 3.2e-28;
Matches 513; Conservative 0; Mismatches 531; Indels 21; Gaps 3;

QY 260 TTGACGAGAGGAAATCCAGATGAACCTTTGAGAAATGGAAGCCCTAGGGCTTT 319
Db 3657964 TGGACGAGAGCGCGGTTCGCCGAAGAGCGCTAGTGGCGCTCAATTCCTCCGGTTTCA 3657905
QY 320 TTGGCTGCAAGTCCCGAAGAATATGGTGGCTGGCTTCTCCAAACACCATGTACTCAA 379
Db 3657904 ACGCGTCCACATTCGGAGGAGTACGCGGTGAGGTGCGGCTAGCTACCTGCA 3657845
QY 380 GACTAGGGAGATCATCAGCATGGATGGTCCATCATCTGTGACCTGGCAGCGCACCGAG 439
Db 3657844 TCGTGATCGAAGAGGTGGCGCGCTCGACGGCTCGCTGTGATCCCGCGGTCAACA 3657785
QY 440 CTATTGGCTCAAGGGGATCATCTTGGCTGAGGAGTGGAGGAGGAGGAGGAGGAGGAG 499
Db 3657784 AGCTGGGACCATGGGCTGATCTGGCGGCTCGAGGAGTGAAGAGGAGGTGCTGC 3657725
QY 500 CTAACTGGCTCGGGGAGCACATTCAGCTTCTGCTCAGCGAGCGACCGAGTGGGA 559
Db 3657724 CGGCGTTGGCGCGGAGGCGGATGGCTTCTATGATTCAGTGAACGCGCGCGGCA 3657665
QY 560 GCGATGAGCTCAATCCGGAGCAGACCCACATAGTGAAGACAGAACGACTACATCC 619
Db 3657664 GTAGCGGCTCCATCGGCGACGCGGCGC-----AAGCGGATGGGATCACATGGATTC 3657611
QY 620 TCAATGGCTCAAGGTCTGATTAATGAGGAGTGGCAATATTTTACTGTGTTG 679
Db 3657610 TCAAGCGGCAAGTCTGGATCACCAGCGGCAAGTCACTGTTACCGGTATGG 3657551
QY 680 CAAGACTGAGTCTGTTGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 739
Db 3657550 CGGTGACCGA-----TCCCGACCGGCGCGCAACGGCATCTCGCGGTTCATGG 3657503

QY 740 TAGAAGAGACTTTGTGGAGTCACTAATATGGAACCCGGAAGATAAATTAGCATTCGGG 799
Db 3657502 TGCACAAGCAGCAGCAGGAGGTTCCACCGTCCGGTCCGAAACGCAAGCTCGGATCAAGG 3657443
QY 800 GCTCCAACACTTTGTGAAGTCCATTTTGAACACCAAGATCTGTGGAAACATCCTTTG 859
Db 3657442 GGTCAACGACCAACGAGCTGTACTTCGAGAACTGCGCATCCCGCGGATCGCATCATCG 3657383
QY 860 GAGAGTCCGAGATGGTTTAAAGTGGCCATCAACATCTCAACAGCGCGCGTTTCAGCA 919
Db 3657382 GTGAGCCGGTACCGGTTTCAAGACCGCGCTCGCCACGCTTGACACACACCCCTCC 3657323
QY 920 TGGGAGCGTCTGGCTGGCTGCTCAAGAGATTTGATGAATGATGCTGATGATGCTGCT 979
Db 3657322 TTGGCGCGCAGGCGCTGGGTATCGCCAGGCGCGCTGGATCTGCCATCGCTACACCA 3657263
QY 980 GCACAAGGAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTTGATTCAGGAGAAATTTG 1039
Db 3657262 AGGACCGCAAGCAGTTTCGTTGAGTGCATGATGATTTTTCAGGCGCTGAGTTCATGCT 3657203
QY 1040 CACTGATGCTCAGAAAGCTTACGTCATGGAGAGTATGACCTTACCTACAGCAGGATGC 1099
Db 3657202 CCGACATGCGATGAAGGTGGAGGCGCTCGGCTGATGTTTACTCCCGCGCGC---C 3657146
QY 1100 TGGACCAACCTGGCTTTCCCGACTCTCCATCGAGGAGCGCATGGTGAAGGTGTTACGT 1159
Db 3657145 CGCGCAAGCGGTGAGCGGATCTGGGCTTTCATTTCCGCGCGCTTCCGAAGTCTCGCT 3657086
QY 1160 CCGAGCGCGCTGGCAGTGTGTGAGTGGCGCTGCACATCTCGGGGCTTTGGGCTACA 1219
Db 3657085 CCGAGTCCGATGGAGTCAACCGATCGCTACAACTGTTGGCGCGCGCGGCTACA 3657026
QY 1220 CAAGGACTTCCGTACAGCGCATACTGCGTGACACCGCGCTCTCTCTCTCTCTCTCT 1279
Db 3657025 CCACCGACTTCCCGTTCGAGCGGTTTATGTCGCGACGCCAAGATCACCGATCTATGAG 3656966
QY 1280 GAACCAATGAGATTCCTCGGATGTACATCGCCCTGACGGGTCTGC 1324
Db 3656965 GCACCAATCAGATTCAGCGGCTAGTGTGTCGGGCGGCTACTGC 3656921

RESULT 7

US-09-221-017B-824/c
; Sequence 824, Application US/09221017B
; Patent No. 644799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIORITY DATA:
; APPLICATION NUMBER: PPI182
; FILING DATE: 31-DEC-1997
; PRIORITY DATA:
; APPLICATION NUMBER: PPI546
; FILING DATE: 30-JAN-1998
; PRIORITY DATA:

Db 686 GTAGCAAGACTTTCATCAGCCAGCCCGGACGCGGACTTCGCCATCGTCTTCGCCGTCA 745
Qy 686 CTGAGTCGTTGATTTCTGATGGATCACTGAAAGACAAATACAGCATTCATAGTAGAAA 745
Db 746 CCGACAGCTACGAGCACAACGGCGGACGCGGAGCGGAGCGGCTTCCTGGTGACA 805
Qy 746 GAGACTTTGGTGAGTCACTAAATGGGAAACCCGGAAGATAAATAGGCATTCGGGGGTCCA 805
Db 806 AGGCGACCGCGGATGACCGTGCAGCGGCGGCGGAAATCGTGAGCAACCGTGGCTACC 865
Qy 806 ACACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGGAAACATCCTTTGGAGAGG 865
Db 866 ACACCTACGAGATCTCTTCGAGCATTCGCGGCTGCGGCTCCAGGTCCTCGGGGAGG 925
Qy 866 TCGGAGATGGGTTAAAGTGGCCATGAACATCCTCAACAGCGCGGCTTCAGCATGGCA 925
Db 926 TCGGCAAGGCTGGGAAGTGCCCAACGCTGCTACCGCGGACGCGGTGATGTCGCGG 985
Qy 926 GCTCGTGGTGGCTGCTCAAGAGATTTGATGAAATGATGCTGAGTAGCCTTGACAA 985
Db 986 CCAACTGTGTGCGCCAGGCGGCGGCGGCTCGACCTGCTGCTGCGCTGGCGGCGGATC 1045
Qy 986 GGAACAGCTTAAACAAGAGCTCAGTGAATTTGGATGATTCAGGAGAAATTTGCACTGA 1045
Db 1046 GCAAGCAGTTCGCCAGCGCATCGGAGCTACCGAGGCGTTCCTTCAAGCTCGCGGACA 1105
Qy 1046 TGCTCAGAGGCTTACGTATGAGAGTATGACCTACCTCACAGCAGGAGTCTGGAGC 1105
Db 1106 TGCCACGAGATCCGTGCGCGGAGCTGATGACCTGCACACCGCTGGAAGATGAGC 1165
Qy 1106 AACCTGGCTTCCGACTGCTCATCGAGCAGCCATGTTGAAGTGTTCAGCTCCGAGG 1165
Db 1166 AGGCAACCATGACCGACGG-----CGAGCGCGCATGGCCAACTGTTCCGAGCGAGA 1219
Qy 1166 CCCTCGGAGTGTGTGAGTGAGGCGCTGCAGATCCTCGGGGCTTGGGCTACACAGGG 1225
Db 1220 CCCTCGGAGGTCGCGGAGAGCGGTGCAGATCTTCGGCGCATGGGCGCTGATGGATG 1279
Qy 1226 ACTATCCGACGAGCGCATCTGCGTACACCGCATCTCTCTCATCTTCGAGGGAACCA 1285
Db 1280 AAGGACCGTTCGAGCGCATCTGGCGCAACGCGGATGCAACGGATCTGGAGGGCACTT 1339
Qy 1286 ATGAGATTCCTCGGATGATACATCGCCCTGACGGGTCTGCAGC 1327
Db 1340 CGGAATCCAGCGGACATCGTTTCCCGGAACCTGCTGCGGC 1381

RESULT 9

US-09-252-991A-16061/c
; Sequence 16061, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16061
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16061

Query Match 7.2%; Score 134; DB 4; Length 1665;
Best Local Similarity 48.8%; Pred. No. 2.4e-27;
Matches 430; Conservative 0; Mismatches 440; Indels 12; Gaps 2;

Qy 446 GCCTCAAGGGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCAATACTTGCCTAAAC 505
Db 1364 GGCCCTCGAAGATCTCTATGGCTGCACCGGAGAACAGACTTCGGCGACTACCTGTGGCCCT 1305
Qy 506 TGGCGTCCGGGGAGCACATTTGCAGCCTTCTGCCTCACGGAGCAGGACGAGTGGGAGGATG 565
Db 1304 GCGTGCAGGCGGAAAGACCGATTTGCTTCGCCCTCACCGAGCCGGCGGCTCCGACG 1245
Qy 566 CAGCCTCAATCCGGAGCAGAGCCACACTAGTGAAGAACAGAGCACTACATCCTCAATG 625
Db 1244 CCAATTCGATCAAGACCCGCGGTGCGGACGCGCAGC-----CTTCGTGATCAACG 1191
Qy 626 GCTCCAAGGTCTGATTTACTAATGAGGAGTGGCCAAATATTTTACTGTGTTTGGCAAGA 685
Db 1190 GTAGCAAGCACTTCATCAGCAGCGGCGGACCGGACTTCGGCCATCGTCTTCGCCGTCA 1131
Qy 686 CTGAGTCTGTTGATCTGATGATAGTGAAGACAAATACAGCATTCATAGTAGAAA 745
Db 1130 CCGACAGCTACGAGCACAACGGCGGCAACGCGTACCGGCTTCCTTGGTGACA 1071
Qy 746 GAGACTTTGGTGAGTCACTAATGGAAACCCGAGATTAATTTAGGCATTCGGGGGTCCA 805
Db 1070 AGGCGACCGCGGATGACCGTGCAGCGGCGGCGGAAATTCGTCAGCAACCGTGGCTACC 1011
Qy 806 ACACCTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACAACTCTTGGAGAGG 865
Db 1010 ACACCTACGAGATCTTCTTCGAGGATTCGCGGTGCGGCTCCCAAGTGTCTCGGCGAG 951
Qy 866 TCGGAGATGGGTTAAGTGGCCATGAACATCCTCAACAGCGCGGTTTTCCTTCAAGCTCGCGGACA 925
Db 950 TCGGCAAGGCTGGGAAGTGGCAACCGCTGCTCACCGCGGACGGTGTGCTGCGCG 891
Qy 926 GCTCGTGGCTGGCTGCTCAAGAGATTTGATTTGAATGACTGCTGAGTAGGCTTCGACAA 985
Db 890 CCAACTGTGTCGGCGAGCGGCGGCGGCTCGACTGCTGCGCTGGGCGCGCGATC 831
Qy 986 GGAACAGCTTAAACAGAGGCTCAGTGAATTTGGATTTGATTCAGGAGAAATTTGCACTCA 1045
Db 830 GCAAGCAGTTCGCGCAGCGCATCGGAGCTACAGGCGGCTTTCCTTCAAGCTCGCGGACA 771
Qy 1046 TGGCTCAGAGGCTTACGTATGAGAGATGATGACCTACCTACAGCAGGAGTGTGGAGC 1105
Db 770 TGGCCACGAGATCGGTGCGCGGAGCTGATGACCTGTCACACCGCTGGAAGATGGACC 711
Qy 1106 AACCTGGCTTCCCGACTGCTCATGAGGAGCAGCATGTTGAAGTGTTCAGCTCCGAGG 1165
Db 710 AGGCGACCATGACCGACGG-----CGAGCGGCGATGCCAAGCTGTTCCGCCAGCGAGA 657
Qy 1166 CCGCTCGCAGTGTGTGAGTGAGGCGCTGCAGATCCTCGGGGCTTGGGCTACACAAGG 1225
Db 656 CCCTCGGCAAGGTCCCGAGCAAGCGGTGCAGATCTTCGGCGGATGGGCTGTGGATG 597
Qy 1226 ACTATCCGTACAGCGCATCTGCGTGACACCGCGATCCTCTCATCTTCGAGGGAACCA 1285
Db 596 AAGGACCGGTGCGAGCGCATCTGCGCGCAACGCGCGGATGCAACGGATCTGGGAGGCACTT 537
Qy 1286 ATGAGATCTCCCGGATGTACATCGCCCTGACGGGTCTGCAGC 1327
Db 536 CGGAATCCAGCGGACATCGTTTCCCGGGAACCTGCTGCGGC 495

RESULT 10

US-08-311-731A-121
; Sequence 121, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:

ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
 STREET: 600 ATLANTIC AVENUE
 CITY: BOSTON
 STATE: MASSACHUSETTS
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/311,731A
 FILING DATE:
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: GATES, EDWARD R.
 REGISTRATION NUMBER: 31,616
 REFERENCE/DOCKET NUMBER: C0044/7125
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/720-3500
 TELEFAX: 617/720-2441
 INFORMATION FOR SEQ ID NO: 121:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 33312 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: MYCOBACTERIUM LEPRAE
 US-08-311-731A-121

Query Match 7.1%; Score 132.6; DB 4; Length 33312;
 Best Local Similarity 47.4%; Pred. No. 2.7e-26;
 Matches 510; Conservative 0; Mismatches 544; Indels 21; Gaps 3;

260 TTGACGAGGAGGAAATCCAGATGAAATTTTGAGAGAAATGAAAGAGCCCTAGGGCTTT 319
 24500 TGACACAGCGTGTCTGATTTCCCGAGGAGCGGTGCGAGCCCTGATGATCATGTTCA 24559
 320 TTGGGCTGCAAGTCCAGAGAAATATGGTGGCTGGGCTTCTCCACACCATGATCTCAA 379
 24560 ACGTATCCAGTTCCTCCGAGGAGTATGGTGTGAGGCTGGGATTCGGTAGCGGCTTGA 24619
 380 GACTAGGGAGATCATCAGCATGGATGGGTCCATCTGTCAGTGGACCTGGCAGCGCACAGG 439
 24620 TTGTGATCGAAGAGTGGCGGTGTCTGATGCTTCTGATGCTTCTGATGCTTCTGATGCTT 24679
 440 CTATTGGCTTCAAGGGGATCATCTTGGCTGGCAGTGGAGAGAGAAAGCCAAATATCTTC 499
 24680 AGCTTGACCATCGGACTCATCTCGCGGTTGCGAAGAGCTCAAGAAACAGGTTCTGC 24739
 500 CTAAGTGGCTGGGGAGCACATGTCAGCCCTTCTGCTCAGGAGCGCACGATGGGA 559
 24740 CATCTGTGGCTGGGAGGGGCGGATGCGCTTATGCTTAACTGAGCGCGAAGCGCGCA 24799
 560 GCGATGAGCTCAATCCGAGGAGAGCCACACTAAGTGAAGACAAAGACACTACATCC 619
 24800 GTACGCTGGCTGATGCGGACCCGGGCCAAA-----GCTAGGGGATGATGATTC 24853
 620 TCAATGGCTCCAAGGCTTGGATTAATGAGGAGTGGCCAATATTTTACTGTGTTG 679
 24854 TCAATGGCTTCAAGTGTGATTAACCAACGCTGCAAGTGCAGCTGTTACACGCTTATGG 24913
 580 CAAAGACTGAGTGTGATTCATGATGATGAGTGAAGACAAATACAGCATTCATAG 739
 24914 CGGTGACGATTCGCGAAGAGGGGCCA-----ACGGCATCTCGGCGTTATCG 24961
 740 TAGAAGAGACTTTGTGTGAGTCACTAATGGGAAACCCGGAAGATATAATAGGCAATTCGGG 799

24962 TGCACAAGACGATGAGGAGTTCAGATTGGCCCCGAAAGAAAGAGTTCGGATCAAGG 25021
 800 GCTCCAACACTTGTGAAGTCCATTTTGAAGAACACCAAGATACCTGTGGAAACATCTCTTG 859
 25022 GGTACCAACACCGAAGCTTACTTCGATAAATGCGATCCCGGTGATGCGATCATTTG 25081
 860 GAGAGTCGGAGATGGGTTTAAGGTGCCATGAACATCTCAACACCGCGCGGTTTCAGCA 919
 25082 GTGAGCCCGGTACTGGCTTTAAGACACAGCGCTAGCCACGTTGGATCATCACGCGTCCACGA 25141
 920 TGGGACGCTGCTGGCTGGCTGCTCAAGAGATTGATTGAAATGACTGCTGAGTAGCGCT 979
 25142 TTGGTGCCCAAGCGGTGGGCATTTGCCAGGGGCGGTTGGACGCTGCCATCGTTTATACCA 25201
 980 GCACAAGAAACAGTTTAAACAGAGCTCAGTGAATTTGATTTGATTCAGGAGAAATTTG 1039
 25202 AGGACCGCAAGCAATTCGCGGAGTGCATTTAGCACTTTCCAGTCCATTGATTCATGCTCG 25261
 1040 CACTGATGCTCAGAAGGCTTACGTCATGAGAGATGACCTACCTCAGACAGGATGC 1099
 25262 CCGACATGGCATGAAAGTGAGGCTGCACGGTTAATTGTCTA---CGTGGCGGTGCC 25318
 1100 TGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGCAGCCATGTTGAAGGTTTTCAGCT 1159
 25319 GTGCTGAACCGGTGAGCCGGATCTGGGCTTTATTTACAGGGGCTGAAATGCTTTGCTT 25378
 1160 CCGAGGCGCTGGCAGTGTGAGTGAGGCGCTGCAGATCTCTCGGGGGCTTGGGCTACA 1219
 25379 CCGACATTCGATGAGGTCAACACCGAGCTGTGCAATTTGTTGGCGCGCGGCTACA 25438
 1220 CAAGGAGCTATCGGTACGAGCGCATCTGCTGACACCCGCACTCTCTCATCTTCGAGG 1279
 25439 CTTCCGACTTCCCGCTGAGCGGTTTCATGCGCAGCGCAAGATCACACAGATCTATGAGG 25498
 1280 GAACCAATGAGATTCTCCGGATGTACATGCGCTGACGGGTCTGACAGCATGCCGG 1334
 25499 GGCACCAATCAGATTCAGCGTGTGATGTCGCGGGGCTGCTGCGCTGATCCGG 25553

RESULT 11

US-09-252-991A-14526
 ; Sequence 14526, Application US/09252991A
 ; Patent No. 6551795

GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 14526

; LENGTH: 1260

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-14526

Query Match 6.6%; Score 123.2; DB 4; Length 1260;

Best Local Similarity 47.8%; Pred. No. 2.1e-24;

Matches 504; Conservative 0; Mismatches 523; Indels 27; Gaps 4;

262 GACCAGGAAGGGGAAAATCCCATGATGAACCTTTGGAGAAATTTGAAGAGCCCTTAGGCTTTT 321
 220 GACCGCGAGCACCGTTTCCCGCGGAGGCAATCCGCGAGATGGCCGACCTTCGGCTTCCTC 279

322 GGCTGCAAGTCCAGAGAAATATGTGGCTTCTCCAAACACCATGACTCAA-- 379
 280 GGCATGCTGTGTCGCCGAGGAATGGGCGCGCGCAGACCGGGCACCTTGGCCTAGCCATG 339

QY	1099	CTGGACCAACCTGGCTTTCCCGAGCTGCTCCATCGAGCGACGCCATGGTGAAGGTGTTTCAGC	1158
Db	225	CGCGAA-----CGCGGCTTCGCTGCTGCTACCGAAGCCTCGATGGCCACGCTGTTGCGC	172
QY	1159	TCGAGAGCGCCCTGGCAGTGTGTGAGTGAAGCGCTGCAGATCCTCGGGGCTTGGGCTAC	1218
Db	171	TCGAGATGCCCCGAGGAAGTTGCTCCGCGGCGATCCAGACCCCTCGGTGGCTATGGCTAC	112
QY	1219	ACAAGGAGTACTCCGTACGAGCGCATACTCGGTGACACCGCGATCCTCCTCATCTTCGAG	1278
Db	111	CTCAAGGACTTCGCGGTGGAACGATCTATCGGGAGCTGCGGGTCTGCCAGATCTACGAG	52
QY	1279	GGACCAATCAGATTCTCCGGATGTACATCGCCC	1312
Db	51	GGCACCAGCGACGTGCAGCGCCTGGTGATCGCGC	18

RESULT 13

US-09-648-004-5

; Sequence 5, Application US/09648004

; Patent No. 6498242

; GENERAL INFORMATION:

; APPLICANT: CHEN, QIONG

; APPLICANT: THOMAS, STUART

; APPLICANT: NAGARAJAN, VASANTHA

; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIIC ACID AND

; TITLE OF INVENTION: INTERMEDIATES

; FILE REFERENCE: CL-1341-A

; CURRENT APPLICATION NUMBER: US/09/648,004

; CURRENT FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/252,553

; PRIOR FILING DATE: 1999-02-19

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 5

; LENGTH: 1155

; TYPE: DNA

; ORGANISM: Acinetobacter sp.

US-09-648-004-5

Db 1103 AATCCAACAGGTCATTATTGCC 1126

RESULT 14

US-09-648-004-27

; Sequence 27, Application US/09648004

; Patent No. 6498242

; GENERAL INFORMATION:

; APPLICANT: CHEN, QIONG

; APPLICANT: THOMAS, STUART

; APPLICANT: NAGARAJAN, VASANTHA

; TITLE OF INVENTION: BIOLOGICAL METHOD FOR THE PRODUCTION OF ADIPIC ACID AND

; TITLE OF INVENTION: INTERMEDIATES

; FILE REFERENCE: CL-1341-A

; CURRENT APPLICATION NUMBER: US/09/648,004

; CURRENT FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/252,553

; PRIOR FILING DATE: 1999-02-19

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: Microsoft office 97

; SEQ ID NO 27

; LENGTH: 17417

; TYPE: DNA

; ORGANISM: Acinetobacter sp.

US-09-648-004-27

```

1880 AAATCCAGCTGAATCTGTCAGCAAAATGAAGAACTGGGTCTTTTGGTCTCACCATT 1939
QY 335 CAGAAGAAATATGGTGGCTGGCTGCTCCCA---ACACCATGTACTCAAGACTAGGGGAGA 391
Db 1940 CTGAGGAATATGAGGGTCTGGCTGACCATGGAGGAAGAGTTTACATTTGCAATTTGAAC 1999
QY 392 TCATCAGCATGGATGGGTCTCATCTGTGACCTGGCAGCGCACCAGGCTATTGGCCCTCA 451
Db 2000 TGGGACGTACCTCTCCTGCTTCCGTTCACTGATCGGCACATAACAATGGGATCGGTTTCAT 2059
QY 452 AGGGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTGCCTTAAACTGGCGT 511
Db 2060 CAGCTTAATATTGATGGCTGCGAGAGCAGAAACAGATATTTTGGCCACGCTGGCAA 2119
QY 512 CCGGGAGCAGCATGTGACGCTTGTGCTCAGGAGCGCAGCAGCTGGAGGATGACGCT 571
Db 2120 GTGGTGAATATTGGTTCAATCTGTTAACTGAACCTGATTCGGGTTTCAGATGCTGCT 2179
QY 572 CAATCCGGAGCAGGCACACTAAGTGAAGCAGAAAGCAAGCACTACATCCTCAATGGCTCCA 631
Db 2180 CTTT-----AAAAACACAGCGGTGAAGATGGTGATCATATTACATTTAATGGCACTA 2233
QY 632 AGGTCTGGATTACTAATGGAGGACTGGCCAAATATTTTACTGTGTTTGAAGAGACTGAGG 691
Db 2234 AGGCTTACATCAACAATGACCGCATGCGGGTGTCTTTACTGTCTATGTCATGGCAGCTAC ----- 2288
QY 692 TCGTTGATTCTGATGATCAGTGAAGACAAAATCACAGCATTCATAGTAGAAGAGACT 751
Db 2289 ----CAGTACCGAAATTAAGGTACAGGTGGAATTTTCAGCCTTTTATCGTGGAGCAATAA 2344
QY 752 TTGCTGGATCACTAATGGAAACCGAAGATAAATTAGGCATTTCCGGGCTCCCAACACTT 811
Db 2345 CTCCTGGTATTTCTTGGTAAACGTGATAAGAGATGGCCAAAGGTGCACATACCT 2404
QY 812 GTCAAGTCCATTTTGAACACACAGATACCTG---TGAAACATCTCTTGGAGAGTGG 868
Db 2405 GTGATGTGATTTTGAACATGTGCTGATTTCTGTCATCTGCATCTCATTTGGTGTGTAAG 2464
QY 869 GAGATGGGTTTAAAGTGGCCATGAACATCCTCAACAGCGCGCGTTTCAGCATGGCAGCG 928
Db 2465 GTGTAGGTTTAAACGTGCAATGAAGTACTTGAATAAGCGCGTATTCATATGCTGCAT 2524
QY 929 TCGTGGCTGGCTGCTCAAGAGATTCATGAATGACTGCTGAGTACGCTGCAACAAGGA 988
Db 2525 TAAGTGTAGTGTGCTGCTGCGGTATGCTGGAAGATTTCCCTACAATATGCGGTTGAGCGCA 2584
QY 989 AACAGTTTAAAGAGGCTCAGTGAATTTGGATTTGATTCAGGAGAAATTTGCATGATGG 1048
Db 2585 AACAGTTTGGTCAAGCGATTGCGAACTTCCAGTTGATTCAGGATGATGATGCGGATTTCTA 2644
QY 1049 CTCAGAGGCTTACGTCATGGAGATGATGACCTACCTCACAGCAGGGATGCTGGACCAAC 1108
Db 2645 AAGCTGAATTTACGACGCAAAATATGATGATGATGATGCTGCGCGACTTCGTGA----- 2699
QY 1109 CTGGCTTTCCGACTGCTCCATCGAGGAGCGCATGTTGAAGGTGTTAGCTCCGAGCGCG 1168
Db 2700 -TGCTGGACAGATGTCAGCAGGAAGCATCTGTGCGCAAGATGTTTGGCCACTGAAATGT 2758
QY 1169 CTGCGAGTGTGTAGTGAGGCGCTGCAATCCTCGGGGCTGCGGCTGACACAAGGACT 1228
Db 2759 GTGGCGGTGTCGACAGATCGTGGCGTACAGATCCATGTTGGTGGCGGTTATATCATGTAAT 2818
QY 1229 ATCCGTACGAGCGCATCTGCTGACACCCCGCATCTCCTCATCTTCGAGGGAACCAATG 1288
Db 2819 ATGCTATTGAGCGTTTATACCGTGTATGATGATGATGATGATGATGATGATGATGATGAT 1312
QY 1289 AGATTCTCCGGATGATATCGCCC 1312
Db 2879 AAATCCACAGGTCATATTGCCC 2902

```

RESULT 15

```

US-09-328-352-2254
; Sequence 2254, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 2254
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-2254

```

```

Query Match 5.7%; Score 105.8; DB 4; Length 1170;
Best Local Similarity 47.3%; Pred. No. 1.4e-19;
Matches 502; Conservative 0; Mismatches 532; Indels 27; Gaps 5;
QY 257 AAATTCACAGGAGGAAATCCAGATGAAACTTTGGAGAAATTTGAAGAGCCTTAGGCG 316
Db 101 AAGTGTAGAGAAACACAAATTCAGATGACATTTGTCAGCAATGCTGAGCTGGGTT 160
QY 317 TTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGGCTTCTTCCAAACACCATGTACT 376
Db 161 TATTTGGCTTCACAAATCCCGAAGAAATATGTTGGCTTAGGCATCCTATGGAAGAGGA 220
QY 377 CAAGACT---AGGGAGATCATCAGCATGGATGGTGGTCCATCCTGTCAGCTGGCAGGCG 433
Db 221 TCAGAGTTGCAATTTGAACCTTGACAACTTCACCGGCTTTTCGTTCTTTAATTTGTTACCA 280
QY 434 ACCAGGCTATTGGCTTCAAGGGGATCATCTTGGCTGGCACTGAGGAGCAGAGAAAGCCAAAT 493
Db 281 ATAACGGTATTGGCTCAAGTGCATTTTGGATGATGGCAGAGAGAGCAAAAGCAGAAAT 340
QY 494 ACTTGGCTTAACTGGCTCCGGGAGCAGCATTTGAGGCTTCTGCTCAGGAGGAGCGCA 553
Db 341 ATTTACCACTGTTACGCAAGTGGCGAAATATTGTTTTCATTTTAAACGAGCTGAAT 400
QY 554 GTGGAGGATGCAAGCTCAATCCGGAGCAGAGCCACATAGTGAAGAGAGAGAGCACT 613
Db 401 CTGTTCTGATGCTGCTCTTTTAAACACAGTGGCGTA-----AAAGATGGCGATTTT 454
QY 614 ACATCTCAATGGCTCCAAGTCTGATGATTAATAAGGAGCTGGCCAAATTTTACTG 673
Db 455 ATGATTAAACGGGCAACAAAGCTTTTATACCAATGCAACCCCATGCTGCAACTTTTACC 514
QY 674 TGTTCGAAAGACTGAGTGTGATTTCTGATGATGATGATGATGATGATGATGATGATGAT 733
Db 515 TAATGGCTCGGAC-----GAATCCTGAATTAAGGGTGGGTTGCGGATTTTCAGCTT 565
QY 734 TCATAGTAGAAGAGACTTTTGGTGGATCACTAATGGGAAACCCGAGATGAAATTTAGGCA 793
Db 566 TTTTATTCGAGGCGCAATACACCGGCAATACATTTAGGCAAAATTTGACCAAAATTTGGGC 625
QY 794 TTCGGGCTCCAACTTGTGAAGTCCATTTTGAACCAACCAAGATACCTG---TGGAAA 850
Db 626 AAAGGGTTCACATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 685
QY 851 ACATCTTGGAGAGGCTGGAGATGGGTTTAAAGTGGCACTGATGATGATGATGATGATGAT 910
Db 686 TGATTGGTGGCGTAGAAGTGTGCGTTTAAACAGCGCATGAAAGTGTGATGATGATGATGAT 745
QY 911 GGTTCAGCATGGCAGCGCTGCTGGCTGGCTGCTCAAGAGATTTGATTTGAATGACTGCTG 970
Db 746 GTTTGCACATTTGGTGGCTATAGCTGGGTGTTGCCGAGCGCATGTTTAAACGAGCGCAT 805
QY 971 AGTACGCTGCACAGGAAACAGTTTAAACAGAGGCTCAGTGAATTTGATTTGATTTGATTT 1030
Db 806 ACTATCGGATTGAGCGTTAAGCAGTTTGGTGCAGCCTATTGCGAACTTTTCAGCTTATTCAAG 865

```

1031	QY	AGAAATTTGCATGTATGGCTCAGAAAGGTTAGTCATGGAGAGTATGACCTACCTCACAG	1090
866	Db	CCATGCTTCCGACCTCAAAGCGGAGATTATGCAAGCTAAATGATGTGCTCTAGATGACG	925
1091	QY	CAGGAGATCTGACACCAACTGSCCTTTCCGAGCTGCTCCATCGAGCAGCATGTTGAAGG	1150
926	Db	CTCGCGCGCGTGTAATGGCG-----AAACAATTAGTACAGAAGCATCTTCGCAAAAA	979
1151	QY	TGTTTCAGTCCGAGGCGCGCTGGCAGTGTGTGAGTGAGCGGTGAGATCTCTCGGGGCT	1210
980	Db	TGTTTTGCGACTGAAATGTTCGGACAGCTGTGTGTCGACCGTGTGTACAGATTCACGTTGGG	1039
1211	QY	TGGGCTACACAAGGAGACTATCGTAGGAGCGCATACTGCTGCACACCCGATCCTCTCA	1270
1040	Db	CGGCTATATCATGTGATGCGATTGACCGTTTTATCGAGATGTGCGTTTTGTTCCGGT	1099
1271	QY	TCTTGAGGGAACCAATGATTTCCGGATGTACATCGCC	1311
1100	Db	TATATGAAGGAACCCGCAAGTTTCAACAACCTATTATTGCC	1140

Search completed: September 6, 2003, 12:58:37
Job time : 120.074 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2003, 23:10:50 ; Search time 3919.86 Seconds
(without alignments)
11551.252 Million cell updates/sec

Title: us-09-945-326-3

Perfect score: 1863

Sequence: 1 atgagcggctgcgggctctt.....accctctggacagcatgc 1863

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1374.4	73.8	2432	11	AK075984
2	1371.8	73.6	2446	11	AK031820
3	1178.4	63.3	1914	11	BC041572
4	1074.6	57.7	2906	11	AK028571

5	937.8	50.3	1201	13	BX458691
6	936.6	50.3	1071	12	BM808820
7	910.2	48.9	1201	9	AL539220
8	859.6	46.1	1010	12	BM552756
9	845.4	45.4	1079	12	BM561207
10	819.6	44.0	969	9	AL549834
11	810.6	43.5	884	14	CA488487
12	810	43.5	922	13	BUS28453
13	800.4	43.0	851	13	BUI76038
14	798.8	42.9	888	13	BQ960770
15	798.4	42.9	890	13	BQ938039
16	798	42.8	1201	9	AL560956
17	797.6	42.8	937	13	BQ279135
18	797.4	42.8	1199	9	AL556077
19	796.2	42.7	885	13	BUI91988
20	796.2	42.7	905	13	BUS41780
21	792.4	42.5	1023	13	BQ072180
22	780.2	41.9	891	13	BX328269
23	777.2	41.7	914	9	AL524997
24	772.4	41.5	851	10	BG757623
25	772.2	41.4	875	13	BQ882994
26	772	41.4	1023	12	BM809053
27	755	40.5	831	12	BM524078
28	750	40.3	867	10	BG674212
29	747	40.1	980	12	BQ054246
30	741.2	39.8	798	12	BI762564
31	739.4	39.7	784	12	BI829609
32	734.4	39.4	857	9	AL557535
33	732.4	39.3	746	12	BI113771
34	732.4	39.3	799	12	BI914477
35	724.2	38.9	893	12	BI772602
36	718.8	38.6	799	12	BI70667
37	718.2	38.6	965	10	BG109786
38	717.2	38.5	973	10	BE792118
39	717	38.5	845	13	BUI53157
40	716	38.4	748	12	BG764035
41	707	37.9	913	10	BE795758
42	704.6	37.8	771	9	AUI42885
43	704	37.8	745	12	BG760511
44	703.6	37.8	942	13	BX350634
45	702.8	37.7	832	12	BI911835

ALIGNMENTS

RESULT 1
AK075984
LOCUS
DEFINITION
AK075984. 2432 bp mRNA linear HTC 07-DEC-2002
Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2600017P15 product:VERY-LONG-CHAIN ACYL-COA
DEHYDROGENASE VLCAD homolog [Homo sapiens], full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

AK075984.1 GI:26344893
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,N., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagakawa, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, F., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2432)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanaoka, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

1. .2432
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2600017P15"
/db_xref="MGI:1895339"
/db_xref="taxon:10090"
/clone="2600017P15"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days embryo"
29. .1906
/note="unnamed protein product; VERY-LONG-CHAIN ACYL-COA
DEHYDROGENASE VLCAD homolog [Homo sapiens] (SPTRIAL56011,
evidence: FASTV, 85.6%ID, 99.6%length, match=1869)
putative"
/codon_start=1
/protein_id="BAC36096.1"
/db_xref="GI:26344874"
/translation="MSGCVLLSRGATAAAARASRVLRRTARRRPLHTLSQCSFA
KELFLGNIEQGVFPPEVDSQHELSEINQFVGLPEKFFTEEDSRKIDQEGKIPVDTL
BKLSGLFGLQVPEYGGGLSNTMYARLGELISLDASITVTLAAHQAIGLKGILV
CNEOKAKYLPKLSSEHIAFCLTEPASGDAASIQTRATLSEDKYKFLNGSKWVI
TNGGLANFTFAKTEVDSGSKTKMTAFIVERFGSITNGKEDKLGIRGNTCE
VHFENTRVPEVGLGVGFKVAMNINLSNFRSGMSAVAGMKLIELTAEVACTRK
QFNRLSFGLIQEKFALMAQKAYVNESMAYLTSGMDQPPGPPSEIKAMVKVFSSE
AAQCVSALQIIGSGYMKDYPERMLRDARILLIFETNEILRLFTALTQLQHAGR
ILTSRIKELKSGNVTVMETIGRKLRLDRTVDLGLTDLGVHPSLGSANKLEEN
VHFQRTVETLLRFGKIVEQLVLRVANILINLYGMTAVLSRASRIRIGLRNHD
HEVLNMFCEVAYFQNLFSQLDKNAPENLDEQIKKVSQILEKRAYICAHPLDRA
S"
BASE COUNT 640 a 564 c 652 g 576 t
ORIGIN
Query Match 73.8%; Score 1374.4; DB 11; Length 2432;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 1546; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

QY 32 CGGCTGCGGCTCGTCTGCGGGTCTGTGGTCTCTACCGGGAACCGGGGCTACTGC 91
DB 72 CGGCTGCGGCGGAGCTTCCCGGGTCTGAGGAATTCACGGCGGCGCCCGCTGC 131
QY 92 GCACACCGCGCTGTACGAGCTTTCGCCAAAGAGCTTTTCTAGCAAAATCAAGAAGA 151
DB 132 ACACAGCGCTGACAGAGCTGCTGCTGCCAGGAGCTCTTCTGGCAACATCAGCAGA 191
QY 152 AAGAGCTTTTCCCATTTCCAGAGTTAGCAAGATGAACATTAATGAATCAATCAGTTCT 211
DB 192 AAGAGCTTTTCCCATTTCCAGAGTTCAGTCAACATGAACATTAATGAATCAATCAGTTCT 251
QY 212 TGGGACCGCTGGAATAATTTCTTCACTGAAGAGTGGACTCCCGAAAATTTGACAGAG 271
DB 252 TGGGACCGCTGGAATAATTTCTTCACTGAAGAGTGGACTCTCGAAAATTTGACAGAG 311
QY 272 GGAATAATCCAGATGAACATTTTGGAGAAATTTGAAGAGCTTAGGCTTTTGGGCTCAAG 331
DB 312 GGAATAATCCAGTTCACCTTAGAGAGTTGAAGAGCTTGGGACTTTTGGCATACAGG 371
QY 332 TCCGAGAAGATATGTGGCTGGCTGGCTTCTTCAACACCATGTACTCAAGACTAGGGAGA 391
DB 372 TCCGAGAAGATATGTGGCTGGCTGGCTTCTTCAACACCATGTACTTGGCTGGGGAGA 431
QY 392 TCATCAGCATGGATGGTCCATCCTGTGACCTTGGCAGCGCACAGGCTATTGGGCTCA 451
DB 432 TCATCAGCTTGGATGGCTTATCCTGCTGACCTTGGCAGCGCACAGGCTATTGGGCTCA 491
QY 452 AGGGGATCATTTTGGCTGGCTAGGAGCAGAAACCAAAATTAATTTGGCTAAACTGGCT 511
DB 492 AGGGGATCATTTTGGCTGGCTAGGAGCAGAAACCAAAATTAATTTGGCTAAACTGGCT 551
QY 512 CCGGGAGCAGCATTTGAGCTTCTGCTTACGAGGAGCAGGCTGGGAGGATGAGCAGCT 571

Db 552 CTGGGGAACACATCGCCGCTTCTGCTTAACAGAGCCAGCGAGTGGAGTGATGCTCGT 611
 QY 572 CAATCCGAGCAGACACACTAAGTGAAGACAAGACCACTACATCTCAATGGCTCCA 631
 Db 612 CCATCCAGACCAAGACTACGTTAAGTGAAGATAAGAAGTACTTCACTCAATGGCTCCA 671
 QY 632 AGGTCTGGATTACTAATGGAGGACTGCCCAATATTTTACTGTGTTTGCAAAAGACTGAGG 691
 Db 672 AGSTTTGGATCACCACCAATGGAGGACTGCCCAATATTTTACTGTGTTTGCAAAACCCGAGG 731
 QY 692 TCGTTGATCTTGATGGATCAGTGAAGACAAAATCAGCAGATTCATAGTAGAAGAGACT 751
 Db 732 TGGTCGATCTGATGGTTCGAAACACAGACAAAATGACCGCATTCATAGTAGAAGAGACT 791
 QY 752 TTGGTGGAGTCACTAATGGGAAACCCGAAGATAAATTAGGCATTCGGGGCTCCAAACACTT 811
 Db 792 TCGCGGAATCACTAATGGGAACCTGAACATAAATTAGGCATTCGAGGCTCCACACAT 851
 QY 812 GTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACACATCTTGGAGAGGTCGGAG 871
 Db 852 GTGAGGTCCATTTTGAATAATACAGAGTGCCTGTGGAATAATGCTTGGAGAAGTTGGAG 911
 QY 872 ATGGTTTAAGGTGGCCATCAACATCTCAACAGCGCGGTTTCAGCATGGGAGCGTGC 931
 Db 912 GGGCTTTAAGGTAGCCATGAACATCTGAACAGTGGCGATTCAGCATGGGAGTGCCTG 971
 QY 932 TGGCTGGCTGCTCAAGAGATTGATTGAATGACTGCTGAGTACGCTGCACAGCAAAAC 991
 Db 972 TGGCGGGATGCTGAAGAGCTGATCGAATGACTGCTGAGTATGCTGTACGAGGAAC 1031
 QY 992 AGTTTAAACAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACTGATGGCTC 1051
 Db 1032 AGTTCAACAGGAATCTCAGTGAATTTGGTTGATCCAGGAAAGTTTGGCTCATGGCTC 1091
 QY 1052 AGAAGCTTACGTCATGGAGATGATGACCTACCTACAGCAGGATGCTGGACCAACCTG 1111
 Db 1092 AGAAGCATACGTAATGGAGAGTATGCTTACCTCACTTCAGGATGCTAGACCAACCA 1151
 QY 1112 GCTTCCCGACTGCTCCATCGAGGAGCCATGGTGAAGTGTTCAGCTCCGAGCGGCT 1171
 Db 1152 GATTTCCCGACTGCTATCGAGGAGCCATGGTGAAGTGTTCAGCTCCGAGCGGCT 1211
 QY 1172 GGCAGTGTGATGAGGCGCTGCAGATCCTCGGGGCTTGGCTTACACAGGAGTATC 1231
 Db 1212 GGCAGTGTGAGGAGGCTTCAGATCCTTGGGGCTTCAGGCTACATGAAGAGTACC 1271
 QY 1232 CGTACAGGCGATCTGCTGACACCCCGCATCCTCCTCATCTTCAGGGAACCAATGAGA 1291
 Db 1272 CCTATGAGCGCATGCTGCTGATGCGCCGATCTCCTAATCTTCAGGGAACCAATGAGA 1331
 QY 1292 TTCTCCGGATGTACATCGCCCTGACGGTCTGCAGCATCGCCGCGCATCTGACTACCA 1351
 Db 1332 TTCTCCGGTGTTCATTTGCCCTGACAGGCTTCAGCATGCTGAGCGCATCTGACTCGA 1391
 QY 1352 GGATCATGAGCTTAAACAGGCCAAGTGAAGCAGTATGATACCTGTTGGCGGAGGC 1411
 Db 1392 GGATCAAGAGCTTAAAGTGGCAATGTGACCACATGTCAGGAGCATTTGGTCGGAAC 1451
 QY 1412 TTCCGGACTTCTGGCGCAACTGTGGACCTGGGGCTGACAGGCAACCATGGATGTGC 1471
 Db 1452 TTCCGGATCTCTGGCGCGCAACTGTGGACCTGGGGCTTAACAGGCGATCTTGGTGTGTC 1511
 QY 1472 ACCCCGACTTCTGGCGACAGTGCACAAGTTTGGAGCAACACCTACTGCTTCGGCGGA 1531
 Db 1512 ACCCCAGCTTGGCGACAGTGCACAAGCTTGGAGAAATGTCCATTACTTTTGGCGGA 1571
 QY 1532 CGGTGGAGACTGCTGCTCGCTTTGGCAAGACCATCATGAGGAGGAGCTGTTACTGA 1591
 Db 1572 CTGTTGACCCCTGTTACTCGCTTTTGGAAAGAACATTTAGAGGAACAGCTGTTGTA 1631
 QY 1592 AGCGGGTGGCAACATCTCATCAACCTGATGCGATGACGGCGCTGCTGTCGGGCGCA 1651
 Db 1632 AGCGGGTAGCCAACATCTCATCAACTTGTATGCGATGCTGCTGTCGAGAGGCA 1691

QY 1652 GCGCTCCATCCGCTTGGGCTCCGCAACACAGCAGGTTCTCTTTGGCAACACCT 1711
 Db 1692 GCGCTCCATCAGAAATTTGGCTCAGGAATCATGATCAGGAGTCTCTTTGGCAACATCT 1751
 QY 1712 TCTGCTGGAAGCTTACTTTGCGAATCTCTTACGCTCTCTCAGCTGGACAAGTATGCTC 1771
 Db 1752 TCTGCTGGAAGCTTACTTTCCAGAACTCTCTTACGCTCTCTCAGCTGGACAAGAGCTC 1811
 QY 1772 CAGAAACCTAGATGAGCAGATTAGAAAGTGTCCAGCAGATCTCTTGAAGAGCGAGCT 1831
 Db 1812 CAGAAACCTGATGAGCAGATTAGAAAGTGTCTCGCAGATCTCTTGAAGAGCGAGCT 1871
 QY 1832 ATATCTGTGCTCCACCTCTCGGACAGCATGC 1863
 Db 1872 ATATCTGTGCTCCACCTCTCGGACAGCATGC 1903

RESULT 2
 AK031820
 LOCUS
 DEFINITION
 Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
 enriched library, clone:6330407H09 product:VERY-LONG-CHAIN ACYL-COA
 DEHYDROGENASE VLCD homolog [Homo sapiens], full insert sequence.
 ACCESSION
 AK031820
 VERSION
 GI:26327642
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 PUBMED
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 PUBMED
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multiplexed capillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 PUBMED
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schirml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anon, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Carlhold, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2446)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resesc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers

1. 2446

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB.6330407H09"

/db_xref="taxon:10090"

/clone="6330407H09"

/sex="male"

/tissue_type="medulla oblongata"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

42. .1919

/note="unnamed protein product; VERY-LONG-CHAIN ACYL-COA DEHYDROGENASE VLCAD homolog [Homo sapiens] (SPTF:AA156011, evidence: FASTA, 85.6%ID, 99.6%length, match=1869) putative"

/codon_start=1

/protein_id="BAC27565.1"

/db_xref="GI:26327643"

/translation="MSGCVLLSRGATAGAAARASRVLRRETBARRRLHPTSLQSCSFA KELTGNTEQGVFPPEVPSQHSLEINQFVPLEKFTTEVDSSRKIDQPKRIPVDL ELKLSGLGFLQVPEEYGGLSNMYARLGEIIISLDASITVTLAAHQALGKILIV NEQKALYKPLKSLGGHIAFLCTEPASGDAASIQTFRATLSDEKPYFLINGSKWVI TNGKLANFTVFAKTEVVDGSKTDKMTAFIVERDFGINTGPKPEKLRGNSNICE VHFENTRPVENLVEGVGGFKVAMNINLSGRFSMSAVAGMLKLLTELPAEYACTRK QFNRLNLSFGLQEKFALMAQKAYVMESMAYLTGMLDOPGPDPCSLEAMVKVFSSE .AAQCVSEALQILGGSGYMKDYPYERMLQRDLRIILIFEGTNEILRLRIALTGLQHAGR ILTSRIKELKSGNVTVMETIGRKLRSLSRGTVDLGLTGLGVVHPHSLGDSANKLEEN

TITLE	JOURNAL	MDLINE	PUBLISHED	REFERENCE	AUTHORS
Query Match	73.6%;	Score 1371.8;	DB 11;	Length 2446;	
Best Local Similarity	84.3%;	Pred. No. 0;	Mismatches 287;	Indels 0;	Gaps 0;
Matches 1544;	Conservative	0;			
BASE COUNT	641 a	567 c	660 g	578 t	
ORIGIN					
QY	33	GGCTGGCTGCTGCTGCTGCGGGGTCTGTGTCTCTCTCGCGGAACCGCGCTACTCGG	92		
Db	86				
QY	93	CACGAGCGCGCTGTACGAGCTTTCGCCAAAGAGAGCTTTTCTAGGCAAAATCAACAAGAA	152		
Db	146	CACGAGCTGCGAGAGCTGCTCGTTCGCCAAGAGAGCTTCTCTGGGCAACATCGAGCAAA	205		
QY	153	AGAAGTTTCCCATTTCCAGAAAGTTAGCAAGCAAGCAACTTAATGAATCAATCACTGCTTT	212		
Db	206	AGGAGTTTCCCATTTCCAGAGGTGAGTCAACATGAATCTAGTGAATCAATCACTGCTGCT	265		
QY	213	GGGACCGCTGGAAAAATTTCTTCACTGAAGAGGTGAGCTCCCGAAAAATTTGACGAGAGG	272		
Db	266	GGGACCGCTGGAAAAATTTCTTCACTGAAGAGGTGAGCTCTCGAAAAATTTGACGAGAGG	325		
QY	273	GAAATCCAGATGAACACTTTGAGAAATTTGAGAGCCTAGGCTTTTGGGCTGCAAGT	332		
Db	326	AAAAATCCAGATTTGACACCTTAGAGAGTGAAGAGCTGGGACTTTTGGCATACAGT	385		
QY	333	CCGAGAGAAATATGGTGGCTTCTCCCAACACCATGTACTCAAGACTAGGGAGAT	392		
Db	386	CCGAGAGAAATATGGTGGCTTCTCCCAACACCATGTACTCGCTCGGCGAGAT	445		
QY	393	CATCAGCATGGATGGGTCCATCACTGTGACCTTGGCAGCGCAGCCAGGCTATTGGGCTCAA	452		
Db	446	CATCAGCTGGATGGCTCTATCAGGTGACCTTGGCAGCAGCAGCAGGCTATCGGCTCAA	505		
QY	453	GGGATCAGTCTGGCTGGCAGTGAAGAGCAAGAAAGCAAAATCTTCCTTAAACTTGGCGTC	512		
Db	506	GGGATATCTTGGTGGCAATGAGAAACAGAGGCAAGTATCTGCCCAACTGTCGTC	565		
QY	513	CGGGAGCAGATTCGAGCCTTCTGCTCAGCGAGCCAGCCAGTGGGAGGATGACACCTC	572		
Db	566	TGGGGAACACATCGCAGCCTTCTGCTTAACAGAGCCAGCCAGTGGGAGTGTGCTCGTC	625		
QY	573	AATCCGGAGCAGACCCACACTAAGTGAAGCAAGAGCACTACATCTCAATGGCTCCAA	632		
Db	626	CATCCAGACAGAGCTACGTAAAGTGAAGATAAGAAGTACTTCTCACTCAATGGCTCCAA	685		
QY	633	GGTCTGGATTACTAATGGAGGACTGGCCAAATATTTTACTTGTGTTTGCAGAAAGTGAAGT	692		
Db	686	GGTTGGATCACCAGTGGAGGACTGGCCAAATATTTTACTTGTGTTTGCAGAAAGTGAAGT	745		
QY	693	CGTTGATTCGATGGATCACTGAAGCAAGAAATACAGACATTCATCTAGTGAAGAGACTT	752		
Db	746	GGTCTGATCTGTGGTTTGGAAACACAGACAAATAGACGCATTCATAGTAGAAGAGACTT	805		
QY	753	TGTTGGAGTCACTAATGGGAAACCCGAAAGATAAATTTAGGCATTTGGGGCTTCCCAACTTG	812		
Db	806	CGCGGAAATCACTAATGGGAAACCTGAAGATAAATTTAGGCATTCGAGGCTTCCCAACTG	865		
QY	813	TGAAGTCCATTTTGAAGAACACCAAGATACCTGTGGGAAACATCTCTTGGAGAGTTCGAGA	872		
Db	866	TGAGTCCATTTTGAAGAAATACCAGAGTGCCTGTGGAAATATCTCTTGGAGAGTTCGAGG	925		
QY	873	TGGTTTAAAGTGCCATGAACCTCCTCAACAGCGCGCTTCCAGCATGGGCGCTGCTGCT	932		
Db	926	GGGCTTAAAGTAGCCATGAACATCTTCAACAGTGGCGGATTCAGCATGGCAGTGCCTG	985		
QY	933	GGCTGGGCTGCTCAAGAGATTGATTGAATGAAATGACTGTGAGTACGCCCTGCACAGGAACA	992		
Db	986	GGCGGGGATGCTGAAGAGGCTGATCGAACTGACTGCTGAGTATGCTTGCCTGTACGAGGAACA	1045		

```

QY 993 GTTTAAACAAGAGCTCAGTGAATTTGGATTGATTCAGAGAAATTTGCACTGATGCTCA 1052
Db 1046 GTTCAACAGGAATCTCAGTGAATTTGGTTGATTCAGAGAAATTTGCGCTATGCTCA 1105
QY 1053 GAAGGCTTACGTCATGAGAGATGATGATCTACCTCAGAGAGGATGCTGACCAACCTGG 1112
Db 1106 GAAGGATACGTAATGAGAGATGATGCTTACCTCAGTTCAGGATGCTAGACCAACGAG 1165
QY 1113 CTTTCCCGACTGCTCCATCGAGGACCCATGTTGAAGTGTTCAGTCCGAGGCGCCCTG 1172
Db 1166 ATTTCCCGACTGCTCCTATCGAGGACCCATGTTGAAGTGTTCAGTTCGAGGCTGCTG 1225
QY 1173 GCAGTGTGAGTGAAGGCTGCGATCCTCGGGGCTTGGCTACACAGGACTATCC 1232
Db 1226 GCAGTGTGAGGAGGCTGCTGAGATCCTTGGGGCTCAGGCTACATGAAGGACTACC 1285
QY 1233 GTACGAGCCATACCTGCTGACACCCGATCTCTCTCATCTTCGAGGGAACCAATGAGAT 1292
Db 1286 CTATGAGCCATGCTGCTGATGCGCGCATCTCTCTATCTTTGAGGAACCAATGAGAT 1345
QY 1293 TCTCCGATGTATACCTGCGGCTGACGGTCTGAGCATGCGCGCATCTCTGACTACAG 1352
Db 1346 TCTTGGTGTTCATTTGCCCTCAGAGGCTGAGCATGCTGAGCGATCTCTGACTCAG 1405
QY 1353 GATCCTAGCTTAACAGGCGCAAGTGAACAGTCACTGATGATCCGTTGGCGGAGGCT 1412
Db 1406 GATCAAGAGCTTAACAGTGGCAATGTGACCAAGTCACTGAGAGCGATTGTCGGAAC 1465
QY 1413 TCGGACTCCCTGGCGCAAGTGTGAGTGGCTGAGGCTGACAGCAACCATGAGTTGTGCA 1472
Db 1466 TCGGATCTCTGCGGCGCAAGTGTGAGTGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 1525
QY 1473 CCCAGTCTTGGGAGTGGCCAAAGTTTGGAGGAAACACTACTGCTCGGCGGAC 1532
Db 1526 CCCAGCTTGGCGAGTGGCCAAAGTTTGGAGGAAATGTCCATTAATTTGGCGGAC 1585
QY 1533 CCGTGGAGACCTGCTGCTCCGCTTTGGCAAGACCATCATGAGGAGCAGTGGTACTGAA 1592
Db 1586 TGTGAGACCTGCTGCTCCGCTTTGGCAAGCAATCATGAGGAGCAGTGGTACTGAA 1645
QY 1593 CGGGTGGCCAAACATCTCATCACTGTATGGCATGACGCGCTGCTGCTCGGCGGAC 1652
Db 1646 CGGGTGGCCAAACATCTCATCACTGTATGGCATGACTGCTGCTGCTGAGGAGCAG 1705
QY 1653 CCGCTCCATCGCATTTGGGCTCGGCAACGACGACGAGTCTCTTTGGCCAAACACTT 1712
Db 1706 CCGCTCCATCGCATTTGGGCTCGGCAACGATCATGATCAGGAGTCTCTTTGGCCAAACACTT 1765
QY 1713 CTGCTGGAAGCTTACTTGCAGATCTCTTCCAGCTCTCTCAGCTGGAACAGTATGCTCC 1772
Db 1766 CTGCTGGAAGCTTACTTCCAGATCTCTTCCAGCTCTCTCAGCTGGAACAGTATGCTCC 1825
QY 1773 AGAAACCTAGATGAGCAGATTAAAGAGTGTCCGAGAGATCTCTTGAAGAGCGAGCTA 1832
Db 1826 AGAAACCTAGATGAGCAGATTAAAGAGTGTCTCGGAGATCTCTTGAAGAGCGAGCTA 1885
QY 1833 TATCTGTCCACCTCTGGACAGCATGC 1863
Db 1886 TATCTGTCCACCTCTGGACAGCATGC 1916

```

RESULT 3

```

BC041572
LOCUS BC041572 1914 bp mRNA linear HTC 02-JAN-2003
DEFINITION Homo sapiens, Similar to NP0002 protein, clone IMAGE:3952086, mRNA.
ACCESSION BC041572
VERSION BC041572.1 GI:27469662
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

1. (bases 1 to 1914)
Strausberg, R.
Direct Submission
Submitted (20-DEC-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

REMARK

COMMENT
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granito, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masilelo, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://Image.llnl.gov>
Series: IRAL Plate: 44 Row: C Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis

This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers
1..1914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3952086"
/tissue_type="Ovary, adenocarcinoma"
/clone_lib="NIH_MGC_9"
/lab_host="DH10B-R"
/note="vector: pOT7"

BASE COUNT 477 a 482 c 522 g 433 t

Query Match

Best Local Similarity 63.3%; Score 1178.4; DB 11; Length 1914;
Matches 1231; Conservative 0; Mismatches 1; Indels 32; Gaps 2;

```

QY 632 AGGTCTGGATTACTAATGGAGGACTGCCAATATTTTACTGTGTTGCAAGACTGAGG 691
Db 133 AGGTCTGGATTACTAATGGAGGACTGCCAATATTTTACTGTGTTGCAAGACTGAGG 192
QY 692 TCGTTGATCTTGATGATCAGTGAAGACAAAATCAGCATTCATAGTAGAAGAGACT 751
Db 193 TCGTTGATCTTGATGATCAGTGAAGACAAAATCAGCATTCATAGTAGAAGAGACT 252
QY 752 TTGCTGGAGTCACTAATGGGAAACCCGAGAGATAAATAGGCATTCGGGGCTCCCAACTT 811
Db 253 TTGCTGGAGTCACTAATGGGAAACCCGAGAGATAAATAGGCATTCGGGGCTCCCAACTT 312
QY 812 GTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACAACTCCTTGGAGAGTGGAG 871
Db 313 GTGAAGTCCATTTTGAACACCAAGATACCTGTGGAACAACTCCTTGGAGAGTGGAG 372
QY 872 ATGGGTTTAAAGTGGCCATGAACATCTCAACAGCGCGCGGTTTACAGCTGGCAGCTCG 931
Db 373 ATGGGTTTAAAGTGGCCATGAACATCTCAACAGCGCGCGGTTTACAGCTGGCAGCTCG 432
QY 932 TGGCTGGGCTGCTCAAGAGATTGATT----GAATGACTGCTGAGTACCCCTGCACAAAG 987

```

433 TGGCTGGGCTGCTCAAGAGATTGATTGGTAGAAATGACTGCTGAGTACGCCCTGCACAAAGG 492
QY 988 AAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTGATT----- 1026
Db 493 AAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTGATTGATTGTTTCTGTTTCTG 552
QY 1027 -----CAGGAGAAATTTGCACTGATCGCTCAGAAAGGCTTACGTCATGGAGATGATGAC 1079
Db 553 AACACTCCAGAGAAATTTGCACTGATCGCTCAGAAAGGCTTACGTCATGGAGATGATGAC 612
QY 1080 CTACCTCAGCAGGAGATGCTGGACCAACTGGCTTTCCCGACTGCTCATCGAGGAGC 1139
Db 613 CTACCTCAGCAGGAGATGCTGGACCAACTGGCTTTCCCGACTGCTCATCGAGGAGC 672
QY 1140 CATGTTGAAGTTTCAGTCCGAGCGCCCTGGCAGTGTGTGAGTGGAGGCTCGAGT 1199
Db 673 CATGTTGAAGTTTCAGTCCGAGCGCCCTGGCAGTGTGTGAGTGGAGGCTCGAGT 732
QY 1200 CCTCGGGGCTTGGGCTACACAGGAGCTATCCGTACGAGCGCATACTGCGTGCACCCG 1259
Db 733 CCTCGGGGCTTGGGCTACACAGGAGCTATCCGTACGAGCGCATACTGCGTGCACCCG 792
QY 1260 CATCTCTCTATCTTCGAGGGAACCAATGAGATTCTCCGATGTACATCGCCCTGACCGG 1319
Db 793 CATCTCTCTATCTTCGAGGGAACCAATGAGATTCTCCGATGTACATCGCCCTGACCGG 852
QY 1320 TCTGCAGCATGCGGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGT 1379
Db 853 TCTGCAGCATGCGGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGT 912
QY 1380 GAGCACAGTCATGATACCGTTGGCGGAGGCTTGGGACTCCCTGGGCGCACTGTGGA 1439
Db 913 GAGCACAGTCATGATACCGTTGGCGGAGGCTTGGGACTCCCTGGGCGCACTGTGGA 972
QY 1440 CTTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGCTTGTGGGAGAGTGCACAA 1499
Db 973 CTTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGCTTGTGGGAGAGTGCACAA 1032
QY 1500 GTTTGAGGAGAACACCTACTGCTTCGCGCGGACCGTGGAGACACTGCTGCTCGCTTTGG 1559
Db 1033 GTTTGAGGAGAACACCTACTGCTTCGCGCGGACCGTGGAGACACTGCTGCTCGCTTTGG 1092
QY 1560 CAAGACCATGAGGAGAGCTGTACTGAGCGGGTGGCCAACTCCTCATCAACCT 1619
Db 1093 CAAGACCATGAGGAGAGCTGTACTGAGCGGGTGGCCAACTCCTCATCAACCT 1152
QY 1620 GTATGGCATGACGGCGCTGCTGCGGGGACCGCTCCATCCGATTTGGGCTCGCA 1679
Db 1153 GTATGGCATGACGGCGCTGCTGCGGGGACCGCTCCATCCGATTTGGGCTCGCA 1212
QY 1680 CCAGCAGCAGAGGTTCTCTTGGCCCAACACCTTCTCGGTGGAGAGCTTACTTGCAGATCT 1739
Db 1213 CCAGCAGCAGAGGTTCTCTTGGCCCAACACCTTCTCGGTGGAGAGCTTACTTGCAGATCT 1272
QY 1740 CTTGAGCTCTTCAGTGCAGAGATGCTCCAGAAACCTAGATGAGCAGATTAGAA 1799
Db 1273 CTTGAGCTCTTCAGTGCAGAGATGCTCCAGAAACCTAGATGAGCAGATTAGAA 1332
QY 1800 AGTGTCCAGCAGATCCTTGAGAGGAGGCTTATCTGTGCCACCCTCTTGACAGGAC 1859
Db 1333 AGTGTCCAGCAGATCCTTGAGAGGAGGCTTATCTGTGCCACCCTCTTGACAGGAC 1392
QY 1860 ATGC 1863
Db 1393 ATGC 1396

RESULT 4

AK028571

LOCUS

DEFINITION Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732402K02 product:VERY-LONG-CHAIN ACYL-CoA DEHYDROGENASE VLCD homolog [Homo sapiens], full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMREFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORSJOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORSTITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORSTITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORSTITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

AK028571
AK028571.1 GI:26080919
HTC: CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamei, M., Lee, N. H., Lyons, P., Nordone, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Sakamoto, N., Sasaki, H.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohtsuki, S.,
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

21085660
11217851

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 2906)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

TITLE
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

1. 2906
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4732402K02"
/db_xref="taxon:10090"
/clone="4732402K02"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
1. 2906
/note="VERY LONG-CHAIN ACYL-COA DEHYDROGENASE VLCAD homolog [Homo sapiens] (SPTRLAAL56011, evidence: FASTY, 85.6% ID, 99.6% length, match=1865)."
BASE COUNT 729 a 673 c 794 g 710 t
ORIGIN

misc_feature

Query Match 57.7%; Score 1074.6; DB 11; Length 2906;
Best Local Similarity 84.9%; Pred. No. 1.4e-283;
Matches 1203; Conservative 0; Mismatches 214; Indels 0; Gaps 0;
QY 447 CCTCAAGGGGATCATCTTGGCTGGACATGAGCAGAGCAAGCAATACTTGCCTAAACT 506
DB 1282 CATTTAGGGGATATCTTGGTCGCAATGAGGAAGCAAGAGGCAAGTATCTGCCAAACT 1341
QY 507 GCGCTCGGGGAGACATATGACGCTTCTGCTCAGGAGCGAGCAGTGGGAGGATGC 566
DB 1342 CTCGCTCGGGGAACACATCGCAGCCTTCTGCTTAACAGAGCCAGTGGAGTGATGC 1401
QY 567 AGCCTCAATCCGAGCAGAGCCACACTAAGTGAAGCAAGAGCACTACATCTCAATGG 626
DB 1402 TCGCTCCATCCAGCAGAGCTACGTTAAGTGAAGTAAGAAGTACTTCACTACTCAATGG 1461
QY 627 CTCGAAGGTCTGGATTACTATGAGGAGTGGCCAAATATTTACTGTTGTTGCAAGAC 686
DB 1462 CTCGAAGGTCTGGATCACCATTGAGGAGTGGCCAAATATTTACTGTTGTTGCAAAAC 1521
QY 687 TCAGTCTCTGATCTGTATGATCAGTGAAGCAAAATCAAGCAATTCATAGTAGAAG 746
DB 1522 CGAGTGTGCTGATCTGTATGTTGGAACAGCAAAATGACCCATTCATAGTAGAAG 1581
QY 747 AGACTTTTGGTGGAGTCACTAATGGAAACCGAAGATAAATAGGCATTCGGGCTCCAA 806
DB 1582 AGACTTCGGCGGATCACTAATGGAAACCGAAGATAAATAGGCATTCGGGCTCCAA 1641
QY 807 CACTTGTGAAGTCCATTTTGAACACCAAGATACCTGTGGAAACATCTTGGAGAGGT 866
DB 1642 CACATGTGAGTCTCATTTTGAATAACAGAGTCCCTGTGGAAATGCTCTTGGAGAGT 1701
QY 867 CGGAGATGGGTTTAAAGTGCCCATGACATCTCTCAACAGCGCGGCTTCAGCATGGGCAG 926
DB 1702 TGGAGGGGGCTTTAAAGTAGCCATGAACATCTCTGAACAGTGGGCGATTTCAGCATGGGCAG 1761

RESULT 5
BX458691
LOCUS
DEFINITION
BX458691 Homo sapiens PLACENTA
5-PRIME, mRNA sequence.
ACCESSION
BX458691
VERSION
BX458691.1 GI:31025045

1201 bp mRNA linear EST 22-MAY-2003
BX458691 Homo sapiens PLACENTA
5-PRIME, mRNA sequence.
BX458691
BX458691.1 GI:31025045

QY 927 CGTCGTGGCTGGCTGCTCAAGAGATTGTTGAATGACTGCTGAGTACGCTGCTGCACAAAG 986
DB 1762 TGCTGTGGCGGGGATGCTGAAGAGCTGATCGAATGACTGCTGAGTATGCTGTCAGGAG 1821
QY 987 GAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTGATTGAGGAGAAATTTCCACTGAT 1046
DB 1822 GAACAGTTTAAACAGAGGCTCAGTGAATTTGGATTGATTGAGGAGAAATTTCCACTGAT 1881
QY 1047 GGTCTCAAGAAGGCTTACGTCATCGAGAGTATGACTACCTACCTACAGCAGGAGTCTGGACCA 1106
DB 1882 GGTCTCAAGAAGGCTTACGTCATCGAGAGTATGACTACCTACCTACAGCAGGAGTCTGGACCA 1941
QY 1107 ACTTGGCTTTCCCGACTGCTTCCATCGAGCAGCAGTATGTTGAAGTGTTCAGCTCCGAGGC 1166
DB 1942 ACCAGGATTTCCCGACTGCTTATCGAGCAGCAGTATGTTGAAGTGTTCAGCTCCGAGGC 2001
QY 1167 CGCCTGGCAGTGTGAGTGAGGCGCTGCAGATCTCTCGGGGCTTGGGGCTACACAAAGGA 1226
DB 2002 TGCTTGGCAGTGTGAGTGAGGCGCTGCAGATCTCTCGGGGCTTGGGGCTACACAAAGGA 2061
QY 1227 CTATCCGTACGAGCGCATACTCGGTGACACCGCATCTCTCTCATCTTCGAGGGAACCAA 1286
DB 2062 CTATCCGTATGAGCGCATCTCGGTGATCGCGCATCTCTCTCATCTTCGAGGGAACCAA 2121
QY 1287 TGAGATTCTCGGATGTACATCGCCCTGACGGGTCTGCAGATGCGCGCCGATCTCTGAC 1346
DB 2122 TGAGATTCTCGGATGTACATCGCCCTGACGGGTCTGCAGATGCGCGCCGATCTCTGAC 2181
QY 1347 TACCAGGATCCATGAGCTTAAACAGGCGCAAGTGAACAGTGAACAGTGAACAGTGAACAG 1406
DB 2182 CTCGAGGATCAAGAGCTTAAAGTGGCAATGTGACCACTGATGAGGAGAGTGGTGG 2241
QY 1407 GAGGCTCGGAGTCCCTCGGCGCAATGTGAGCTTGGGGCTGACAGCAACCACTGAGGT 1466
DB 2242 GAACTTCGGGATCTCTGCGCGCAATGTGAGCTTGGGGCTGACAGGCGATCTGGTGT 2301
QY 1467 TGTGACCCCGAGTCTTGGGAGAGTGGCAACAGTGGAGGAGAGCACTACTGCTTCGG 1526
DB 2302 TGTGACCCCGAGTCTTGGGAGAGTGGCAACAGTGGAGGAGAGCACTACTGCTTCGG 2361
QY 1527 CCGGAGCTGGAGACACTGCTGCTCGGCTTGGCAAGCACTATCATGAGGAGGAGTGGT 1586
DB 2362 CCGGAGTGTGAGACCTGTTTACTCGGCTTGGAAAGCACTATCATGAGGAGAGTGGT 2421
QY 1587 ACTGAAGGGGTGGCAACATCTCTCATCACTGTATGGCATGACGGCGCTGCTGTCGCG 1646
DB 2422 GTTAAAGGGGTAGCCCAACATCTCTCATCACTGTATGGCATGACGCTGCTGTCGAG 2481
QY 1647 GGGCAGCGCTGCCATCCGATTTGGGCTCCGCAACAGCACTATCATGAGGAGGAGTGGT 1706
DB 2482 AGCCAGCGCTGCCATCCGATTTGGGCTCAGGATTCATGATCAGGAGGCTTCTGTTGGCAA 2541
QY 1707 CACTTCTCGGTGGAAGCTTACTTGGCAATCTCTTCAGGCTCTCTCAGCTGGCAAGTA 1766
DB 2542 CATGCTGTGTGGAAGCTTATTTCCAGATCTCTTCAGGCTGCTCAGCTGGCAAGTA 2601
QY 1767 TGCTCCAGAAACCTAGATGAGCAGATTAAGAAAGTGTCCAGCAGATCTCTTGAAGCG 1826
DB 2602 CGCTCCAGAAACCTAGATGAGCAGATTAAGAAAGTGTCTCGGAGATCTCTTGAAGCG 2661
QY 1827 AGCTATATCTGTGCCACCTCTCTGGACAGGATGC 1863
DB 2662 AGCTATATCTGTGCCACCTCTCTGGACAGGATGC 2698

KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	1 (bases 1 to 1201)	
JOURNAL	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.	
COMMENT	Full-length cDNA libraries and normalization	
	Unpublished	
	Contact: Genoscope	
	Genoscope - Centre National de Sequencage	
	BP 191 91006 EVRY cedex - France	
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr	
	Library was constructed by Life Technologies, a division of	
	Invitrogen. This sequence belongs to sequence cluster 2225.f For	
	more information about this cluster, see	
	http://www.genoscope.cns.fr/	
	cgi-bin/cluster.cgi?seq=CS0DE007AF12QP1&cluster=2225.f. Contact :	
	Feng Liang Email : fliang@lifetech.com URL :	
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600	
	Paradise Avenue Genoscope sequence ID : CS0DE007AF12QP1.	
FEATURES	Location/Qualifiers	
source	1..1201	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="CS0DE007YK23"	
	/tissue_type="PLACENTA"	
	/clone_lib="Homo sapiens PLACENTA"	
	/note="Vector: pCMVSPORT6; 1st strand cDNA was primed	
	with a NotI-oligo(dT) primer. Five prime end enriched,	
	double-strand cDNA was digested with Not I and cloned into	
	the Not I and EcoRV sites of the pCMVSPORT 6 vector.	
	Library was not normalized."	
BASE COUNT	290 a 284 c 337 g 240 t	50 others
ORIGIN		
	Query Match	50.3%; Score 937.8; DB 13; Length 1201;
	Best Local Similarity	97.6%; Pred. No. 3.5e-246;
	Matches	983; Conservative 9; Mismatches 11; Indels 4; Gaps 4;
QY	712 GTGAAGACAAATACAGACATTCATAGTAGAAGACATTTGGTGGAGTCACTAATGGG	771
Db		
QY	66 GTGAAGACAAATACAGACATTCATAGTAGAAGACATTTGGTGGAGTCACTAATGGG	125
Db		
QY	772 AAACCCGAAGATAAATAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTGAAC	831
Db		
QY	126 AAACCCGAAGATAAATAGGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTTGAAC	185
Db		
QY	832 ACCAAGATACCTGTGGAACATCCTTTGGAGAGTCGAGATGGGTTTAAAGTGGCCATG	891
Db		
QY	186 ACCAAGATACCTGTGGAACATCCTTTGGAGAGTCGAGATGGGTTTAAAGTGGCCATG	245
Db		
QY	892 AACATCTCAACAGCGCGCTTCAGCATTCGGGGCTCCAGCGCTCGTGGCTGGCTCAAGAG	951
Db		
QY	246 AACATCTCAACAGCGCGCTTCAGCATTCGGGGCTCCAGCGCTCGTGGCTGGCTCAAGAG	305
Db		
QY	952 TTGATTGAATGACTGTGAGTACGCTTCGACAGGAACAGTTTACACAGAGCTCAGT	1011
Db		
QY	306 TTGATTGAATGACTGTGAGTACGCTTCGACAGGAACAGTTTACACAGAGCTCAGT	365
Db		
QY	1012 GAATTGGATTGATTACAGAGAAATTTGCATGTAGTGGCTCAGAGAGGCTTACGTATGGAG	1071
Db		
QY	366 GAATTGGATTGATTACAGAGAAATTTGCATGTAGTGGCTCAGAGAGGCTTACGTATGGAG	425
Db		
QY	1072 AGTATGACCTACCTACAGAGGAGTGTGACCAACCTGGCTTCCGGAGTGTCCATC	1131
Db		
QY	426 AGTATGACCTACCTACAGAGGAGTGTGACCAACCTGGCTTCCGGAGTGTCCATC	485
Db		
QY	1132 GAGGACGACCTGTTAAGTGTTCAGTCCGAGCGCGCTGGCAGTGTGTGAGTGAAGCG	1191
Db		
QY	486 GAGGACGACCTGTTAAGTGTTCAGTCCGAGCGCGCTGGCAGTGTGTGAGTGAAGCG	545
Db		

RESULT 6	
BM808820	
LOCUS	1071 bp mRNA linear EST 05-MAR-2002
DEFINITION	AGENCOURT_6582944 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5471398
5', mRNA sequence.	
BM808820	
VERSION	BM808820.1 GI:19125643
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1071)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCMI979 row: 1 column: 23 High quality sequence stop: 663.
FEATURES	Location/Qualifiers
source	1..1071
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:5471398"
	/tissue_type="amelanotic melanoma, cell line"
	/lab_host="DH10B (phage-resistant)"
	/clone_lib="NIH_MGC_41"

/note="Organ: skin; Vector: pOTB7; Site_1: xhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene), Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT	287 a	253 c	296 g	232 t	3 others
ORIGIN					
Query Match	50.3%; Score 936.6; DB 12; Length 1071;				
Best Local Similarity	96.9%; Pred. No. 7e-246;				
Matches 1017;	Conservative 0; Mismatches 26; Indels 6; Gaps 6;				
QY	1	ATGAGCGGCTGCGGGCTCTTCCTGCCACACAGCGCTGCGGCTGCGGCTGCGGGGTCTG	60		
Db	23	ATGAGCGGCTGCGGGCTCTTCCTTGGCCACACAGCGCTGCGGCTGCGGGGTCTG	82		
QY	61	GTGGTCTCTACCGCGAACCGGGCTACTGCGCACAGCCGCCCTGTAGAGCTTTCGCC	120		
Db	83	GTGGTCTCTACCGCGAACCGGGCTACTGCGCACAGCCGCCCTGTAGAGCTTTCGCC	142		
QY	121	AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGC	180		
Db	143	AAAGAGCTTTTCTAGGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGC	202		
QY	181	CAAGATGAATTAATGAATCAATCAGTTCTTGGACCCGCTGGAAATTTCTTCACTAA	240		
Db	203	CAAGATGAATTAATGAATCAATCAGTTCTTGGACCCGCTGGAAATTTCTTCACTAA	262		
QY	241	GAGTGGACTCCCGAAAATTGACCAGGAAGGAAATCCAGATGAATTTGGAGAA	300		
Db	263	GAGTGGACTCCCGAAAATTGACCAGGAAGGAAATCCAGATGAATTTGGAGAA	322		
QY	301	TTGAAGAGCTAGGCGCTTTTGGGCTGCAAGTCCAGAGATATGTTGGCTGGGCTTC	360		
Db	323	TTGAAGAGCTAGGCGCTTTTGGGCTGCAAGTCCAGAGATATGTTGGCTGGGCTTC	382		
QY	361	TCCAACACATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCCTGTG	420		
Db	383	TCCAACACATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGTCCATCCTGTG	442		
QY	421	ACCTTGGCAGCGCACAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCAGCTGAGAG	480		
Db	443	ACCTTGGCAGCGCACAGGCTATTGGCTCAAGGGATCATCTTGGCTGGCAGCTGAGAG	502		
QY	481	CAGAAAGCAATATCTTGCCTAAACTGGCGTCCGGGAGACATTGGAGCCTTCTGCCTC	540		
Db	503	CAGAAAGCAATATCTTGCCTAAACTGGCGTCCGGGAGACATTGGAGCCTTCTGCCTC	562		
QY	541	ACGAGCGCAGCAGTGGGCGGATGCAGCCTCAATCCGGAGCAGACCACTAAGTGAA	600		
Db	563	ACGAGCGCAGCAGTGGGCGGATGCAGCCTCAATCCGGAGCAGACCACTAAGTGAA	622		
QY	601	GACAAGAGCACTACATCTCAATCGGCTCAAGGTCTGATTAATGAGGAGCTGGCC	660		
Db	623	GACAAGAGCACTACATCTCAATCGGCTCAAGGTCTGATTAATGAGGAGCTGGCC	682		
QY	661	AATATTTTACTGTGTTGCAAGACTGAGGTGCTGATTTCTGATGGATCAGTGAAGAC	720		
Db	683	AATATTTTACTGTGTTGCAAGACTGAGGTGCTGATTTCTGATGGATCAGTGAAGAC	742		
QY	721	AAATACACAGCATTCATAGTAGAAGAGCTTTGGTGGAGTCACTAATGGGAACCGAA	780		
Db	743	CAATACACAGCATTCATAGTAGAAGAGCTTTGGTGGAGTCACTAATGGGAACCGAA	802		
QY	781	GATAAATAGGCAATTC - GGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAGAT	839		
Db	803	GATAAATAGGCAATTC - GGGGCTCCCAACACTTGTGAAGTCCATTTTGAACACCAAGAT	862		
QY	840	ACCTGTGGAACACAT - CCTTGGAGAGGTGGAGAT - GGGTTTAAAGTGGCCATGAACATC	897		

Db	863	ACCTGTGTGAAACATCCCTTTGGAGAGTCCGAGATGGGTTTAAAGTGGCCATGAACATC	922		
QY	898	CTCAACACGGG - CCGGTTTACGAT - GGGCAGGCTGCTGGTGGGCTGCTCAAGAGATTGA	955		
Db	923	CTCAACACGGGCGGTTTCANCATGGGACGCTGCTGGTGGGCTGCTTCAAGAGATTGA	982		
QY	956	TTGAATGACTGCTGAGTACGCTGCACAAGGAAACACTTTTAAACAAGAGGCTCAGTGAAT	1015		
Db	983	TTGAATGACTGCTGAGTACGCTGCACAAGGAAACACTTTTAAACAAGAGGCTCAGTGAAT	1042		
QY	1016	TT-GAATGATTCAGGAGAAATTTGCAT	1043		
Db	1043	TTGGGATGATTCGCGAGAAATTTTGCT	1071		

RESULT 7
AL539220
LOCUS
DEFINITION
AL539220 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
ACCESSION
VERSION
AL539220.2 GI:31263790
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12868235.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2225.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF034A06QPl&cluster=2225.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DF034A06QPl.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF034Y11"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

FEATURES
source

BASE COUNT 320 a 266 c 306 g 244 t 65 others
ORIGIN
Query Match 48.9%; Score 910.2; DB 9; Length 1201;
Best Local Similarity 94.8%; Pred. No. 1.3e-238;
Matches 951; Conservative 27; Mismatches 20; Indels 5; Gaps 4;

QY	1	ATGAGCGGCTGCGG - CCTTCTTCTGCGCACACGCGCTGCGGCTGCGGCTGCGGGGTCT	59		
Db	98	ATGAGCGGCTGCGGNGCTCTTCTGCGCACACGCGCTGCMGCTGCTGCTGCGGGGTCT	157		
QY	60	GGTGTCTCTACCGCAACCGCGGCTACTGCGCACACGCGCTGTACGAGCTTTCGC	119		
Db	158	GGTGTCTCTACCGCAACCGCGGCTACTGCGCACACGCGCTGTACGAGCTTTCGC	217		

RESULT 11	CA488487	884 bp	mrna	linear	EST 14-NOV-2002
LOCUS	AGENCOURT_10808657	MAPcL	Homo sapiens	CDNA clone	IMAGE:6720247 5'
DEFINITION	CA488487	1	GI:24950504		
ACCESSION	CA488487				
VERSION	1				
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	NIH-MGC http://mgc.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: Kristi A. Eglund, Ira Pastan CDNA Library Preparation: Invitrogen Corp DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Cloned distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM14279 row: h column: 07 High quality sequence stop: 713. Location/Qualifiers 1. .884 /organism="Homo sapiens" /mol_type="mrna" /db_xref="taxon:9606" /clone="IMAGE:6720247" /cell_line="2R-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCap" /lab_host="EMDH10B" /clone_lib="MAPcL" /note="pCMV-Sport6; Site_1: EcoRV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligodr. Average				

RESULT	12
BUS28453	
LOCUS	
DEFINITION	BU528453 922 bp mRNA linear EST 13-SEP-2002
	AGENCOURT_10188068 NTH_MCC_101 Homo sapiens cDNA clone

Qy	798	GGGCTCCAACACTTGTGAAGTCCATTTTGAAGAACACCAAGATACCTGTGGAAACATCCT	857
Db	1	GGGCTCCACACTTGTGAANGTCCATTTTGAAGAACACCAAGATACCTGTGGAAAC - TCCT	59
Qy	858	TGGAGAGGTGGGATGGGTTTAAAGTGTGGCCATGAACATCCTCAACAGCGCGCGTTTCAG	917
Db	60	TGGAGAGGTGGGAGATGGGTTTAAAGTGTGGCCATGAACATCCTCAACAGCGCGCGTTTCAG	119
Qy	918	CATGGGCAGCGTCGTGGCTGGGCTGCCTCAAGAGATTGATTCAAATGACTGCTGAGTAGCG	977
Db	120	CATGGGCAGCGTCGTGGCTGGGCTGCCTCAAGAGATTGATTCAAATGACTGCTGAGTAGCG	179
Qy	978	CTGCACAAGGAACAGTTTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATT	1037
Db	180	CTGCACAAGGAACAGTTTAAACAAGAGGCTCAGTGAATTTGGATTGATTCAGAGAAATT	239
Qy	1038	TGCACTCATGGCTCAGAAGGTTACGTCATGGAGAGTATGACCTACCTCACAGCAGGAT	1097
Db	240	TGCACTCATGGCTCAGAAGGTTACGTCATGGAGAGTATGACCTACCTCACAGCAGGAT	299
Qy	1098	GCTGGACCAACCTGGCTTTCCGACTCCTCCATCGAGGCGACCCATGCTGAAGGTTCAG	1157
Db	300	GCTGGACCAACCTGGCTTTCCGACTCCTCCATCGAGGCGACCCATGCTGAAGGTTCAG	359
Qy	1158	CTCCGAGGCGCCCTGGCAGTGTGTGAGTGAAGCGCTGCAGATCCCTGGGGGCTTGGGCTA	1217

```
|||||
360 CTCGAGCGCGCTGCGAGTGTGTAGTGAGGCGCTGCAGATCCTCGGGGCTTGGGCTA 419
QY 1218 CACAAGGACTATCCGTAGAGCGGATAGTGGTGACACCGGATCCTCTCATCTTGA 1277
Db 420 CACAAGGACTATCCGTAGAGCGGATAGTGGTGACACCGGATCCTCTCATCTTGA 479
QY 1278 GGAACCAATGAGATTCGCGATGACATGCCCTGACGGGTCTGCAGCATCCGCGCG 1337
Db 480 GGAACCAATGAGATTCGCGATGACATGCCCTGACGGGTCTGCAGCATCCGCGCG 539
QY 1338 CATCTCTGACTACAGGATCCATGAGCTTAAACAGGCGCAAGTGCAGCATCTGATAC 1397
Db 540 CATCTCTGACTACAGGATCCATGAGCTTAAACAGGCGCAAGTGCAGCATCTGATAC 599
QY 1398 CGTTGCGCGGAGGCTTCGGAGCTCCCTGGGCGGAAGTGTGGACCTGGGCTGACAGG 1457
Db 600 CGTTGCGCGGAGGCTTCGGAGCTCCCTGGGCGGAAGTGTGGACCTGGGCTGACAGG 659
QY 1458 CCATGAGTGTGCACCCAGTCTTCGGGACAGTGCACAAAGTTTGGAGGAGAACCTA 1517
Db 660 CCATGAGTGTGCACCCAGTCTTCGGGACAGTGCACAAAGTTTGGAGGAGAACCTA 719
QY 1518 CTGCTTCGGCGGACCGTGGAGACTGTGCTGCCCTTTGGCAAGACCATCATGAGGA 1577
Db 720 CTGCTTCGGCGGACCGTGGAGACTGTGCTGCCCTTTGGCAAGACCATCATGAGGA 779
QY 1578 GCAGCTGGTACTGAAGC-GGGTGGCAACATCTCATCACTGTATGCATGAAGCGCC 1635
Db 780 GCAGCTGGTACTGAAGC-GGGTGGCAACATCTCATCACTGTATGCATGAAGCGCC 839
QY 1636 GTGCTGTGCGGGGCCAG-CGCTGCCATCCGCTATGGGCTCCGCAACAC--GACCA 1692
Db 840 GTGCTGTGCGGGGCCAGCGCTCCATCCGCGATGGCTCTCGCACCCAGCACCGAG 899
QY 1693 GTTCTCTTGGCCA 1705
Db 900 GTTCTCTTGGCCA 912

RESULT 13
LOCUS BUI76038
DEFINITION AGENCOURT_7782617 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6137085
5', mRNA sequence.
ACCESSION BUI76038
VERSION BUI76038.1 GI:22690022
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 851)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-i@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13450 row: e column: 22
High quality sequence stop: 651.
Location/Qualifiers
1..851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6137085"

FEATURES
source
```

```
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_67"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 234 a 205 c 230 g 181 t 1 others
ORIGIN
Query Match 43.0%; Score 800.4; DB 13; Length 851;
Best Local Similarity 99.4%; Pred. No. 1.7e-208;
Matches 834; Conservative 0; Mismatches 2; Indels 3; Gaps 3;
QY 1 ATGAGCGGCTGGGGCTCTTCTGCGCACCGGCTGCGGCTCGTGCCTCGCGGGTCTG 60
Db 14 ATGAGCGGCTGGGGCTCTTCTGCGCACCGGCTGCGGCTCGTGCCTCGCGGGTCTG 73
QY 61 GTGGTCTTACCGCGAACCGCGGCTTACTGCGCACCGGCTGTACGAGCTTTCGCC 120
Db 74 GTGGTCTTACCGCGAACCGCGGCTTACTGCGCACCGGCTGTACGAGCTTTCGCC 133
QY 121 AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAAGATTTTCCCATTTCCAGAAATAG 180
Db 134 AAAGAGCTTTTCTAGGCAAAATCAAGAAGAAAGATTTTCCCATTTCCAGAAATAG 193
QY 181 CAAGATGAAGTAAATCAAAATCAATCAGTTCTTTGGGACCGGTGGAAATTTCTT 240
Db 194 CAAGATGAAGTAAATCAAAATCAATCAGTTCTTTGGGACCGGTGGAAATTTCTT 253
QY 241 GAGGTGGACTCCGAAAAATTTGACAGGAGGAAAAATCCAGATCAAACTTTGGAGAA 300
Db 254 GAGGTGGACTCCGAAAAATTTGACAGGAGGAAAAATCCAGATCAAACTTTGGAGAA 313
QY 301 TTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGGCTTGGCTTC 360
Db 314 TTGAAGAGCTAGGGCTTTTGGGCTGCAAGTCCCAAGAAATATGTTGGCTTGGCTTC 373
QY 361 TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCTGTG 420
Db 374 TCCAACACCATGTACTCAAGACTAGGGGAGATCATCAGCATGGATGGGTCCATCTGTG 433
QY 421 ACCCTGGCAGCGCACCGAGCTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG 480
Db 434 ACCCTGGCAGCGCACCGAGCTATTGGCTCAAGGGGATCATCTTGGCTGGCACTGAGGAG 493
QY 481 CAGAAAGCCAAATACTTGGCTAAACTGGCGTCCGGGAGCACAATTCAGCCTTCTGCCTC 540
Db 494 CAG-AAGCCAAATACTTGGCTAAACTGGCGTCCGGGAGCACAATTCAGCCTTCTGCCTC 552
QY 541 ACGGAGCCAGCGAGTGGGAGCGATGAGCCTCAATCCGGAGCAGACCCACACTAAGTGAA 600
Db 553 ACGGAGCCAGCGAGTGGGAGCGATGAGCCTCAATCCGGAGCAGACCCACACTAAGTGAA 612
QY 601 GACAAGACCACTACATCCTCAATGGCTCCCAAGGTCTGGATTACTAATGAGGACTGGCC 660
Db 613 GACAAGACCACTACATCCTCAATGGCTCCCAAGGTCTGGATTACTAATGAGGACTGGCC 672
QY 661 AATATTTTACTGTGTGTCAAAAGACTGAGGTCTGTGATTTCTGATGGATGAGTGAAGAC 720
Db 673 AATATTTTACTGTGTGTCAAAAGACTGAGGTCTGTGATTTCTGATGGATGAGTGAAGAC 732
QY 721 AAAATCACAGCATTCATAGTAGAAAGACTTTGGTGGAGTCACTAAT-GGGAAACCCGA 779
Db 733 AAAATCACAGCATTCATAGTAGAAAGACTTTGGTGGAGTCACTAATGGGGAACCCGA 792
QY 780 AGATAAATAGGACTTC-GGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAG 837
Db 793 AGATAAATAGGACTTCGGGGCTCCAAACACTTGTGAAGTCCATTTTGAACACCAAG 851

RESULT 14
BQ960770
```

```

LOCUS      BQ960770      888 bp      mRNA      linear      EST 21-AUG-2002
DEFINITION AGENCOURT_8881366 Lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6201350 5', mRNA sequence.
ACCESSION  BQ960770
VERSION     BQ960770.1 GI:22376248
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE  1 (bases 1 to 888)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13617 row: k column: 15
          High quality sequence stop: 640.

FEATURES   Location/Qualifiers
            1..888
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6201350"
             /sex="male"
             /tissue_type="sciatic nerve"
             /dev_stage="adult, 70 yr"
             /lab_host="DH10B"
             /clone_lib="Lupski_sciatic_nerve"
             /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
             NotI; Site_2: SalI; cDNA made by oligo-dr priming.
             Directionally cloned using the following adapters:
             5'-TCGACCCACGGCTCCG-3' and
             5'-GACTAGTTCATGATCGGAGCGCCGCT(15)-3'. Size selected >
             1 kb for average insert length 1.87 kb. This is a primary
             library, non-amplified. Library constructed by Life
             Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
             College of Medicine) and is available through Life
             Technologies."
BASE COUNT 252 a 194 c 250 g 192 t
ORIGIN

Query Match      42.9%; Score 798.8; DB 13; Length 888;
Best Local Similarity 97.8%; Pred. No. 4.8e-208;
Matches 833; Conservative 0; Mismatches 12; Indels 7; Gaps 2;

QY 194 ATGAATCAATCAGTCTCTGGGACCCGTGGAAAAATTTCTTCACTGAAGAGGTGGACATCCC 253
Db 12 ATGAATCAATCAGTCTCTGGGACCCGTGGAAAAATTTCTTCACTGAAGAGGTGGACATCCC 71
QY 254 GAAAAATTTGACCGAAGGGAATCCAGATGAATCTTTGGAGAAATTTGAAGAGCCCTAG 313
Db 72 GAAAAATTTGACCGAAGGGAATCCAGATGAATCTTTGGAGAAATTTGAAGAGCCCTAG 131
QY 314 GCGTTTTTGGGCTCCAAAGTCCCAAGAAATATGTTGGCCCTGGGCTTCTCCAAACACCATGT 373
Db 132 GCGTTTTTGGGCTCCAAAGTCCCAAGAAATATGTTGGCCCTGGGCTTCTCCAAACACCATGT 191
QY 374 ACTCAAGACTAGGGGAGATCATCAGCATGGATGGTGGTCCATCATCTGACCCCTGGCAGCGC 433
Db 192 ACTCAAGACTAGGGGAGATCATCAGCATGGATGGTGGTCCATCATCTGACCCCTGGCAGCGC 251
QY 434 ACCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGCGACTGAGGACAGAAAGCCAAAT 493
Db 252 ACCAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGCGACTGAGGACAGAAAGCCAAAT 311

```

```

QY 494 ACTTGCCTAACTGGCTCCGGGGAGACACATTTGCAGCCTTCTGCCTCAGGAGCCAGCCA 553
Db 312 ACTTGCCTAACTGGCTCCGGGGAGACACATTTGCAGCCTTCTGCCTCAGGAGCCAGCCA 371
QY 554 GTGGGAGCGATGCAGCCTCAATCCGGAGCAGAGCCACACTAAGTCAAGACAAAGACACT 613
Db 372 GTGGGAGCGATGCAGCCTCAATCCGGAGCAGAGCCACACTAAGTCAAGACAAAGACACT 431
QY 614 ACATCCTCAATGGCTCCCAAGGCTCTGGATTACTAATGGAGGACTGGCCCAATATTTTACTG 673
Db 432 ACATCCTCAATGGCTCCCAAGGCTCTGGATTACTAATGGAGGACTGGCCCAATATTTTACTG 491
QY 674 TGTTTGCAAGACTGAGGTCTGTTGATTTCTGATGATCAGTGAAGACAAAAATCACAGCAT 733
Db 492 TGTTTGCAAGACTGAGGTCTGTTGATTTCTGATGATCAGTGAAGACAAAAATCACAGCAT 551
QY 734 TCATAGTAGAAGAGACTTTTGTGTGAGTCACTAATGGGAAACCCGAAAGATAAATAGGCA 793
Db 552 TCATAGTAGAAGAGACTTTTGTGTGAGTCACTAATGGGAAACCCGAAAGATAAATAGGCA 611
QY 794 TTCGGGGCTCAACACACTTGTGAAGTCCATTTTTGAACACCAACAGATACCTGTGGAAACA 853
Db 612 TTCGGGGCTCAACACACTTGTGAAGTCCATTTTTGAACACCAACAGATACCTGTGGAAACA 671
QY 854 TCCTTTGGAGAGTTCGGAGATGGGTTTAAGGTGGCCATGAACATCCTCAACAGCGCCGCT 913
Db 672 TCCTTTGGAGAGTTCGGAGATGGGTTTAAGGTGGCCATGAACATCCTCAACAGCGCCGCT 731
QY 914 TCAGCATGGGAGCGTCTGGCTGGGCTGCTCAAGAGATTTGATT---GAAATGACTGCT 969
Db 732 TCAGCATGGGAGCGTCTGGCTGGGCTGCTCAAGAGATTTGATTGGTAGAATGACTGCT 791
QY 970 GAGTACCGCTGCACAAAGAAACAGTTTAAACAGAGGCTCAGTGAA---TTTGGATTGATT 1026
Db 792 GAGTACCGCTGCACAAAGAAACAGTTTAAACAGAGGCTCAGTGAA---TTTGGATTGATT 851
QY 1027 CAGGAGAAATTT 1038
Db 852 CCGGGAGAAATTT 863

```

```

RESULT 15
LOCUS    BQ938039
DEFINITION AGENCOURT_8948875 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6469922
          5', mRNA sequence.
ACCESSION BQ938039
VERSION    BQ938039.1 GI:22353517
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens

REFERENCE  1 (bases 1 to 890)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM13999 row: n column: 03
          High quality sequence stop: 611.

FEATURES   Location/Qualifiers
            1..890
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"

```

Tue Sep 9 10:14:39 2003

us-09-945-326-3.rst

```

/clone="IMAGE:6469922"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: Sall; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT      242 a      214 c      248 g      186 t
ORIGIN
Query Match      42.9%; Score 798.4; DB 13; Length 890;
Best Local Similarity 99.5%; Pred. No. 6.2e-208;
Matches 832; Conservative 0; Mismatches 1; Indels 3; Gaps 3;
Qy 1 ATGAGCGGCTGCGGGGCTCTTCCTGCGCACCACGGCTGCGGCTCGTGCCTGCCGGGTCTG 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 55 ATGAGCGGCTGCGGGGCTCTTCCTGCGCACCACGGCTGCGGCTCGTGCCTGCCGGGTCTG 114
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 61 GTGGTCTCTACCGCGAACCGCGGCTACTGCGCACCACGGCGGCTGTAGGAGCTTTTCGCC 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 115 GTGGTCTCTACCGCGAACCGCGGCTACTGCGCACCACGGCGGCTGTAGGAGCTTTTCGCC 174
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 121 AAAGAGCTTTTCTTAGGCAAAATCAAGAGAAAGAGTTTTCCTCATTTCCAGAGTTAGC 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 175 AAAGAGCTTTTCTTAGGCAAAATCAAGAGAAAGAGTTTTCCTCATTTCCAGAGTTAGC 234
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 181 CAAGATCACTTAATGAATCAATCACTTCTTGGACCCGTTGAAATAATCTTCACTGAA 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 235 CAAGATCACTTAATGAATCAATCACTTCTTGGACCCGTTGAAATAATCTTCACTGAA 294
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 241 GAGGTGACTCCCGAAAAATTCAGAGAGGAGAAATCCAGATGAATGAACTTTGGAGAA 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 295 GAGGTGACTCCCGAAAAATTCAGAGAGGAGAAATCCAGATGAATGAACTTTGGAGAA 354
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 301 TTGAAGAGCTTAGGGCTTTTGGGCTCCAGTCCCAGAGAAATATGTTGGCTGGGCTTC 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 355 TTGAAGAGCTTAGGGCTTTTGGGCTCCAGTCCCAGAGAAATATGTTGGCTGGGCTTC 414
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 361 TCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGATGGATGGTCCATCACTGTG 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 415 TCCAAACACCATGTACTCAAGACTAGGGAGATCATCAGATGGATGGTCCATCACTGTG 474
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 421 ACCCTGGAGCGCACAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCAGCTGAGGAG 480
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 475 ACCCTGGAGCGCACAGGCTATTGGGCTCAAGGGGATCATCTTGGCTGGCAGCTGAGGAG 534
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 481 CAGAAAGCAAAATCTTGCCTFAAACTGGCGTCCGGGGAGCACATTGCAGGCTTCTGCCTC 540
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 535 CAGAAAGCAAAATCTTGCCTFAAACTGGCGTCCGGGGAGCACATTGCAGGCTTCTGCCTC 594
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 541 ACGGAGCCAG-CCAGTGGGAGCGATGAGCGCTCAATCCGGAGCAGAGCCACACTAAGTGA 599
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 595 ACGGAGCCAGCCAGTGGGAGCGATGAGCGCTCAATCCGGAGCAGAGCCACACTAAGTGA 654
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 600 AGACAAAGACACTACATCTCAATGGCTCCCAAGTCTGGATTACTAATGAGGACTGGC 659
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 655 AGACAAAGACACTACATCTCAATGGCTCCCAAGTCTGGATTACTAATGAGGACTGGC 714
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 660 CAATATTTTACTGTGTTTGAAGACACTGAGTGTGATTTCTGATGATCAGTGAAGA 719
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 715 CAATATTTTACTGTGTTTGAAGACACTGAGTGTGATTTCTGATGATCAGTGAAGA 774
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 720 CAATATCAGCAATCATAGTAGAAGAGACTTTT-GGTGGAGTCACTAATGGAAACCCG 778
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 775 CAATATCAGCAATCATAGTAGAAGAGACTTTTGGGTGGAGTCACTAATGGAAACCCG 834
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 779 AAGATAAATTAGGCAATTCGGGCT-CCAAACACTTGTGAAGTCCATTTTGAACAC 833
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 835 AAGATAAATTAGGCAATTCGGGCTCCCAACACTTGTGAGTCCATTTTGAACAC 890
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: September 7, 2003, 01:51:44
Job time : 3922.86 secs

Result	No.	Score	Query Match	Length	DB	ID	Description
1	1	1863	100.0	1863	10	US-09-943-336-3	Sequence 3, Appl1
2	2	1863	100.0	2452	10	US-09-943-336-1	Sequence 1, Appl1
3	3	1861.4	99.9	2440	14	US-10-168-274-51	Sequence 51, Appl1
4	4	1084.6	57.1	2080	14	US-10-198-846-11756	Sequence 11756, A
5	5	452.2	24.3	430	11	US-09-918-935-31901	Sequence 21901, A
6	6	451.8	24.3	483	11	US-09-918-935-33707	Sequence 23707, A
7	7	330.6	21.0	451	11	US-09-918-995-11034	Sequence 23707, A
8	8	388.6	20.9	2117	10	US-09-917-800A-1584	Sequence 11034, A
9	9	261.4	14.0	490	11	US-09-918-995-16927	Sequence 1584, Ap
10	10	214.6	11.5	312	9	US-09-822-849A-480	Sequence 16927, A
11	11	165.2	8.9	568	10	US-09-796-632-8175	Sequence 480, App
12	12	165.2	8.9	568	14	US-10-040-862-8175	Sequence 8175, Ap
13	13	159.4	8.6	193	10	US-09-867-701-3329	Sequence 3392, Ap
14	14	131.2	8.1	1158	14	US-10-156-761-1	Sequence 5259, Ap
15	15	131.2	8.1	9025608	14	US-10-156-761-1	Sequence 1, Appl1
16	16	147.6	7.9	1170	14	US-10-156-761-6897	Sequence 6897, Ap

```
Db 121 AAAGAGCTTTTCCCTAGGCAAAATCAAGAAAGAAAGTTTCCCATTTCCAGAGTTAGC 180
QY 181 CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCGCTGGAAATATCTTCACTGAA 240
Db 181 CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCGCTGGAAATATCTTCACTGAA 240
QY 241 GAGGTGGACTCCGGAATAATGACAGAGGAGGAAATCCAGATGAACCTTTGGAGAA 300
Db 241 GAGGTGGACTCCGGAATAATGACAGAGGAGGAAATCCAGATGAACCTTTGGAGAA 300
QY 301 TTCAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGGCTTC 360
Db 301 TTCAAGAGCTAGGGCTTTTGGGCTGCAAGTCCAGAGAAATATGTTGGCTGGCTTC 360
QY 361 TCCAACACCATGTAACAGCTAGGGGAGATCATCAGCATGGATGGGTCCATCACTGTG 420
Db 361 TCCAACACCATGTAACAGCTAGGGGAGATCATCAGCATGGATGGGTCCATCACTGTG 420
QY 421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG 480
Db 421 ACCCTGGCAGCGCACAGGCTATTGGCTCAAGGGGATCATCTTGGCTGGCAGTGGAG 480
QY 481 CAGAAAGCCAAATACCTTAACTGGGCTCGGGGAGCACATTCAGCCTTCTGCCTC 540
Db 481 CAGAAAGCCAAATACCTTAACTGGGCTCGGGGAGCACATTCAGCCTTCTGCCTC 540
QY 541 ACGGAGCCAGCAGTGGGAGCGATGAGCGCTCAATCCGAGCAGAGCCACACTAAGTGAA 600
Db 541 ACGGAGCCAGCAGTGGGAGCGATGAGCGCTCAATCCGAGCAGAGCCACACTAAGTGAA 600
QY 601 GACAAGACGACTACATCCTCAATGGCTCCAGGCTGGAATTAATGGAGGACTGGCC 660
Db 601 GACAAGACGACTACATCCTCAATGGCTCCAGGCTGGAATTAATGGAGGACTGGCC 660
QY 661 AATATTTTACTGTTTGAAGAGCTGAGTGGCTGATCTGATGGATCAGTGAAGAC 720
Db 661 AATATTTTACTGTTTGAAGAGCTGAGTGGCTGATCTGATGGATCAGTGAAGAC 720
QY 721 AATATCAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCGGAA 780
Db 721 AATATCAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAAACCGGAA 780
QY 781 GATAAATTTAGGCAATCGGGCTCCACACTTTGGAAGTCCATTTTGAACACACCAAGATA 840
Db 781 GATAAATTTAGGCAATCGGGCTCCACACTTTGGAAGTCCATTTTGAACACACCAAGATA 840
QY 841 CCTGTGGAACACATCCTTGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC 900
Db 841 CCTGTGGAACACATCCTTGAGAGGTCGGAGATGGGTTTAAAGTGGCCATGAACATCCTC 900
QY 901 AACAGGCGCGGTTACAGATGGGCGGCTGGTGGGCTGCTCAGAGATTGATTGAA 960
Db 901 AACAGGCGCGGTTACAGATGGGCGGCTGGTGGGCTGCTCAGAGATTGATTGAA 960
QY 961 ATGACTGCTCAGTACGCTGCACAGGAACAGTTTAAAGAGGCTCAGTGAATTTGA 1020
Db 961 ATGACTGCTCAGTACGCTGCACAGGAACAGTTTAAAGAGGCTCAGTGAATTTGA 1020
QY 1021 TTGATTACAGAGAAATTTGACATGATGGCTCAGAGGCTTTACGTATGAGAGATGATGACC 1080
Db 1021 TTGATTACAGAGAAATTTGACATGATGGCTCAGAGGCTTTACGTATGAGAGATGATGACC 1080
QY 1081 TACCTCACAGAGGATGCTGGACCACTGGCTTTCCGCTGCTCCATCGAGGAGGCC 1140
Db 1081 TACCTCACAGAGGATGCTGGACCACTGGCTTTCCGCTGCTCCATCGAGGAGGCC 1140
QY 1141 ATGCTGAAGTGTTCAGTCCGAGCGGCTGGCAGTGTGTGAGTGAAGGCGCTGCAGATC 1200
Db 1141 ATGCTGAAGTGTTCAGTCCGAGCGGCTGGCAGTGTGTGAGTGAAGGCGCTGCAGATC 1200
QY 1201 CTGCGGGGCTTGGGCTACACAAGGAGCTATCCGTACAGGCGCATCTGCTGACACCCG 1260
Db 1201 CTGCGGGGCTTGGGCTACACAAGGAGCTATCCGTACAGGCGCATCTGCTGACACCCG 1260
```

```
QY 1261 ATCTCTCTCATCTTCAGGGAACCAATGAGATTCTCCGATGTACATCGCCCTGACGGGT 1320
Db 1261 ATCTCTCTCATCTTCAGGGAACCAATGAGATTCTCCGATGTACATCGCCCTGACGGGT 1320
QY 1321 CTGACGATCGCGGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGTG 1380
Db 1321 CTGACGATCGCGGCGCATCTGACTACAGGATCCATGAGCTTAAACAGGCCAAAGTG 1380
QY 1381 AGCACAGTCAATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1440
Db 1381 AGCACAGTCAATGATACCGTTGGCGGAGGCTTCGGGACTCCCTGGGCGCAACTGTGGAC 1440
QY 1441 CTGGGCTGACAGGCAACCATGAGTTGTCACCCAGTCTTCGGGACAGTGCACAAG 1500
Db 1441 CTGGGCTGACAGGCAACCATGAGTTGTCACCCAGTCTTCGGGACAGTGCACAAG 1500
QY 1501 TTTGAGGAGAAACACTTACTGCTTCGGCGGAGCGCTGGAGACACTGCTGCTCCGCTTTGGC 1560
Db 1501 TTTGAGGAGAAACACTTACTGCTTCGGCGGAGCGCTGGAGACACTGCTGCTCCGCTTTGGC 1560
QY 1561 AGACCATCATGAGGAGGAGCTGTTGTAAGCGGGTGGCCAAACATCTCATCAACCTG 1620
Db 1561 AGACCATCATGAGGAGGAGCTGTTGTAAGCGGGTGGCCAAACATCTCATCAACCTG 1620
QY 1621 TATGGCATGACGGCGCTGCTGCGGGCGGAGCGCTCCATCGCATTTGGGCTCCGCAAC 1680
Db 1621 TATGGCATGACGGCGCTGCTGCGGGCGGAGCGCTCCATCGCATTTGGGCTCCGCAAC 1680
QY 1681 CAGGACGAGAGTTCTTCTTGGCCAAACACTTCTGCTGGAAGCTTACTTTGCAAGATCTC 1740
Db 1681 CAGGACGAGAGTTCTTCTTGGCCAAACACTTCTGCTGGAAGCTTACTTTGCAAGATCTC 1740
QY 1741 TTGAGCTCTCTCAGCTGGACAGATGTTCTCGAAGAAACCTAGATGAGCAGATTAGAAA 1800
Db 1741 TTGAGCTCTCTCAGCTGGACAGATGTTCTCGAAGAAACCTAGATGAGCAGATTAGAAA 1800
QY 1801 GTCTCCAGCAGATCTTGAAGAGCGAGCTATATCTGTGCCACCCCTCTGGACAGACA 1860
Db 1801 GTCTCCAGCAGATCTTGAAGAGCGAGCTATATCTGTGCCACCCCTCTGGACAGACA 1860
QY 1861 TGC 1863
Db 1861 TGC 1863
```

```
RESULT 2
US-09-945-326-1
; Sequence 1, Application US/09945326
; Patent No. US20020127680A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Hunter, John Joseph
; TITLE OF INVENTION: 62112, A NOVEL HUMAN DEHYDROGENASE AND
; FILE REFERENCE: USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/945,326
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/229,831
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)...(1932)
US-09-945-326-1
```

Query Match 100.0%; Score 1863; DB 10; Length 2452;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1863;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------------	--------------	----	------------	----	--------	----	------	----

QY	1	ATGAGCGGCTGCGGGCTCTTCTCGCGACACACGCTGCGGCTGCTGCTGCGGGTCTG	60
Db	67	ATGAGCGGCTGCGGGCTCTTCTCGCGACACACGCTGCGGCTGCTGCTGCGGGTCTG	126
QY	61	GTGCTCTTACC CGGAACCGCGGCTACTGCGCACAGCCGCCCTGTACGAGCTTTTCGCC	120
Db	127	GTGCTCTTACC CGGAACCGCGGCTACTGCGCACAGCCGCCCTGTACGAGCTTTTCGCC	186
QY	121	AAAGAGCTTTTCTTAGGCAAAATCAAGAAAGAAAGTTTTCCTTCCATTTCCAGAAGTTAGC	180
Db	187	AAAGAGCTTTTCTTAGGCAAAATCAAGAAAGAAAGTTTTCCTTCCATTTCCAGAAGTTAGC	246
QY	181	CAAGATGAACCTTAATGAAATCAATCAGTTCTTTGGGACCCGTCGAAAAATCTTCACTGAA	240
Db	247	CAAGATGAACCTTAATGAAATCAATCAGTTCTTTGGGACCCGTCGAAAAATCTTCACTGAA	306
QY	241	GAGGTGACCTCCGGAATAATTGACAGGAAGGAAATCCCAAGATGAAACTTTTGAGAAA	300
Db	307	GAGGTGACCTCCGGAATAATTGACAGGAAGGAAATCCCAAGATGAAACTTTTGAGAAA	366
QY	301	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCACAGAGAAATAGTGGCCTGGGCTTC	360
Db	367	TTGAAGAGCCTAGGGCTTTTGGGCTGCAAGTCCACAGAGAAATAGTGGGCTGGGCTTC	426
QY	361	TCCAACACCATGACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCACTGCTG	420
Db	427	TCCAACACCATGACTCAAGACTAGGGAGATCATCAGCATGGATGGGTCCACTGCTG	486
QY	421	ACCTGCGCAGCGACAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGGCACCTGAGGAG	480
Db	487	ACCTGCGCAGCGACAGGCTATTGGCCTCAAGGGGATCATCTTGGCTGGGCACCTGAGGAG	546
QY	481	CAGAAAGCCAAATACTTGCCTAAACTGGGCTCCGGGAGCACATTCAGGCTTCGCGCTC	540
Db	547	CAGAAAGCCAAATACTTGCCTAAACTGGGCTCCGGGAGCACATTCAGGCTTCGCGCTC	606
QY	541	ACGGAGCCAGCAGTGGGAGCGATGCAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAA	600
Db	607	ACGGAGCCAGCAGTGGGAGCGATGCAGCCTCAATCCGGAGCAGAGCCACACTAAGTGAA	666
QY	601	GACAAGAAGCACTACATCCTCAATGGCTCCAGGCTCTGGATTACTAATGAGGACTGGCC	660
Db	667	GACAAGAAGCACTACATCCTCAATGGCTCCAGGCTCTGGATTACTAATGAGGACTGGCC	726
QY	661	AATATTTTTACTGTGTTGCCAAGACTGAGGCTGGTTGATTCTGATGGATCAGTGAAGAGC	720
Db	727	AATATTTTTACTGTGTTGCCAAGACTGAGGCTGGTTGATTCTGATGGATCAGTGAAGAGC	786
QY	721	AAATCAGAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAACCCGAA	780
Db	787	AAATCAGAGCATTCATAGTAGAAGAGACTTTGGTGGAGTCACTAATGGGAACCCGAA	846
QY	781	GATAAATTAGGCATTCGGGCTCCAAACTGTGAAGTCCATTTTGAACACCAAGATA	840
Db	847	GATAAATTAGGCATTCGGGCTCCAAACTGTGAAGTCCATTTTGAACACCAAGATA	

```

RESULT 3
US-10-168-274-51
: Sequence 51, Application US/10168274
: Publication No. US20030124106A1
: GENERAL INFORMATION:
1. APPLICANT: INCYTE GENOMICS, INC.
2. APPLICANT: YUE, Henry
3. APPLICANT: LAL, Preeti
4. APPLICANT: TANG, Y. Tom
5. APPLICANT: HILLMAN, Jennifer
6. APPLICANT: BAUGHN, Mariah R.
7. APPLICANT: AZIMZAI, Yalda
8. APPLICANT: LU, Dyung Aina M.
9. TITLE OF INVENTION: HUMAN OXIDOREDUCTASE PROTEINS
10. FILE REFERENCE: PF-0754 PCT

```

```

; CURRENT APPLICATION NUMBER: US/10/168,274
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/172,367
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 51
; LENGTH: 2440
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030124106a1 5540437CB1
US-10-168-274-51

```

[illegible]

840	DB	GATAAATTAGGCATTCGGGGCTCCACACTTGTGAAGTCCATTTTGGAAACACCAAGATA	899
841	QY	CCGTGTGAAACATCCTTTGGAGAGTGGAGATGGGTTTAAAGTGGCCATGACATCCTC	900
900	DB	CCTGTGAAACATCCTTTGGAGAGTGGAGATGGGTTTAAAGTGGCCATGAAATCCTC	959
901	QY	AACAGCGCCGGTTACAGCATGGCAGCGTCGTGGCTGGCTCCTCAAGACATGATTGAA	960
960	DB	AACAGCGCCGGTTACAGCATGGCAGCGTCGTGGCTGGCTCCTCAAGACATGATTGAA	1019
961	QY	ATGACTGCTGAGTACGGCTGCACAAGGAAACAGTTTAAACAGAGGCTCAGTGAATTTGGA	1020
1020	DB	ATGACTGCTGAGTACGGCTGCACAAGGAAACAGTTTAAACAGAGGCTCAGTGAATTTGGA	1079
1021	QY	TTGATTTCAGAGAAATTTGCACATGATGGCTCAGAAGCTTACGTCATCGGAGAGTATGACC	1080
1080	DB	TTGATTTCAGAGAAATTTGCACATGATGGCTCAGAAGCTTACGTCATCGGAGAGTATGACC	1139
1081	QY	TACCTACACGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGACGCC	1140
1140	DB	TACCTACACGAGGATGCTGGACCAACCTGGCTTTCCCGACTGCTCCATCGAGGACGCC	1199
1141	QY	ATGTTGAAGTGTTCAGCTCCGAGGCCGCTTGGCAGTGTGTGATGAGGCGCTCCAGATC	1200
1200	DB	ATGTTGAAGTGTTCAGCTCCGAGGCCGCTTGGCAGTGTGTGATGAGGCGCTCCAGATC	1259
1201	QY	CTCGGGGCTTGGCTACACAAGGAGTATCCGTCACGAGCCATCTGCGTGACACCCGC	1260
1260	DB	CTCGGGGCTTGGCTACACAAGGAGTATCCGTCACGAGCCATCTGCGTGACACCCGC	1319
1261	QY	ATCCTCCTCATCTTCGAGGGAAACAAATGAGATTCCTCGGATGTACATCGCCCTGACGGGT	1320
1320	DB	ATCCTCCTCATCTTCGAGGGAAACAAATGAGATTCCTCGGATGTACATCGCCCTGACGGGT	1379
1321	QY	CTGCAGCATGCCGGCCGATCCTGACTACCAGGATCCATGAGCTTAAACAGGCCAAAGTG	1380
1380	DB	CTGCAGCATGCCGGCCGATCCTGACTACCAGGATCCATGAGCTTAAACAGGCCAAAGTG	1439
1381	QY	AGCACAGTCATGGATACCGTTTGGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTGGAC	1440
1440	DB	AGCACAGTCATGGATACCGTTTGGCGGAGGCTTGGGACTCCCTGGGCGCAACTGTGGAC	1499
1441	QY	CTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGTCTTGGGAGAGTGCACAACAG	1500
1500	DB	CTGGGGCTGACAGGCAACCATGGAGTTGTGCACCCAGTCTTGGGAGAGTGCACAACAG	1559
1501	QY	TTTGAGGAGAACACTACTGCTTTCGGCGGAGCCCTGGGAGACACTGCTCCTCGCTTTGGC	1560
1560	DB	TTTGAGGAGAACACTACTGCTTTCGGCGGAGCCCTGGGAGACACTGCTCCTCGCTTTGGC	1619
1561	QY	AAGACCATCATGGAGGACAGTGTGATGAAGGGGTGGCACAACATCCTCATCAACCTG	1620
1620	DB	AAGACCATCATGGAGGACAGTGTGATGAAGGGGTGGCACAACATCCTCATCAACCTG	1679
1621	QY	TATGGCATGAGGGCGTGTGTCGGGCGCAGCCGCTCCATCCGATTTGGGCTCCGCAAC	1680
1680	DB	TATGGCATGAGGGCGTGTGTCGGGCGCAGCCGCTCCATCCGATTTGGGCTCCGCAAC	1739
1681	QY	CACGACCAGAGGTTCTCTTGSCCAACACCTTCTCGGTGGAGGCTTACTTGCAGAACTCTC	1740
1740	DB	CACGACCAGAGGTTCTCTTGSCCAACACCTTCTCGGTGGAGGCTTACTTGCAGAACTCTC	1799
1741	QY	TTTCAGCCCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAAGAA	1800
1800	DB	TTTCAGCCCTCTCAGCTGGACAAAGTATGCTCCAGAAAACCTAGATGAGCAGATTAAAGAA	1859
1801	QY	GTGTCCCGACAGATCCTTTGAGAGGAGGCGCTATATCTGTGCCCAACCTCTTGGACAGGACA	1860
1860	DB	GTGTCCCGACAGATCCTTTGAGAGGAGGCGCTATATCTGTGCCCAACCTCTTGGACAGGACA	1919
1861	QY	TGC 1863	
1920	DB	TGC 1922	


```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21901
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-21901

Query Match      24.3%; Score 452.2; DB 11; Length 490;
Best Local Similarity 99.3%; Pred. No. 1.9e-133;
Matches 454; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1349 CCAGATCCATGAGCTTAAACAGGCCAAAGTGGACACATCATGGATACCGTTGGCCGGA 1408
DB 33 CCAGATCCATGAGCTTAAACAGGCCAAAGTGGACACATCATGGATACCGTTGGCCGGA 92
QY 1409 GGCTTCGGAGTCCCTCGGCGCAAGCTGTGGACCTGGGGCTGACAGCAACCATGGAGTTG 1468
DB 93 GGCTTCGGAGTCCCTCGGCGCAAGCTGTGGACCTGGGGCTGACAGCAACCATGGAGTTG 152
QY 1469 TGCACCCAGTCTTCGGGACAGTGCCCAACAAAGTTTGAGGAGAACACCTACTGCTTGGCC 1528
DB 153 TGCACCCAGTCTTCGGGACAGTGCCCAACAAAGTTTGAGGAGAACACCTACTGCTTGGCC 212
QY 1529 GGACCGTGGAGACACTGCTCTCCGCTTTGGCAAGACCATCATGAGGAGAGCTGGTATG 1588
DB 213 GGACCGTGGAGACACTGCTCTCCGCTTTGGCAAGACCATCATGAGGAGAGCTGGTATG 272
QY 1589 TGAAGCGGTGGCCACATCCTCATCAACCTGTATGGCATGAGCGCGTGTCTGCGCGG 1648
DB 273 TGAAGCGGTGGCCACATCCTCATCAACCTGTATGGCATGAGCGCGTGTCTGCGCGG 332
QY 1649 CCAGCGCTCCATCCGCTATGGCTCCGCAACACACACAGGTTCTCTTGGCCCAACA 1708
DB 333 CCAGCGCTCCATCCGCTATGGCTCCGCAACACACAGGTTCTCTTGGCCCAACA 392
QY 1709 CCTTCTGCGTGGAGCTTACTTCCGAAATCTTTCAGCCCTCTCTCAGCTGGCAAGTATG 1768
DB 393 CCTTCTGCGTGGAGCTTACTTCCGAAATCTTTCAGCCCTCTCTCAGCTGGCAAGTATG 452
QY 1769 CTCAGAAAACCTAGATGACACATTAAGAAAGTGTC 1805
DB 453 CTCAGAAAACCTAGATGACACATTAAGAAAGTGTC 489

RESULT 6
US-09-918-995-23707
; Sequence 23707, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23707
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-23707

Query Match      24.3%; Score 451.8; DB 11; Length 483;
Best Local Similarity 99.6%; Pred. No. 2.5e-133;
Matches 453; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1349 CCAGATCCATGAGCTTAAACAGGCCAAAGTGGACACATCATGGATACCGTTGGCCGGA 1408
DB 29 CCAGATCCATGAGCTTAAACAGGCCAAAGTGGACACATCATGGATACCGTTGGCCGGA 88

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21901
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-21901

Query Match      21.0%; Score 390.6; DB 11; Length 451;
Best Local Similarity 97.8%; Pred. No. 8.3e-114;
Matches 396; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 585 AGCCACACTAAGTGAAGACAAAGACGACATACATCTCAATGGCTCCAAAGTCTGGATTAC 644
DB 47 AGCCACACTAAGTGAAGACATGAAGACATACATCTCAATGGCTCCAAAGTCTGGATTAC 106
QY 645 TAATGGAGGACTGGCCCAATATTTTACTGTGTTTGAAGAGCTGAGGTCTGTTGATCTGA 704
DB 107 TAATGGAGGACTGGCCCAATATTTTACTGTGTTTGAAGAGCTGAGGTCTGTTGATCTGA 166
QY 705 TGGATCAGTGAAGACAAAATACAGCATTCATAGTAGAAGAGAGACTTTTGGTGGAGTCAC 764
DB 167 TGGATCATTGAAGACAAAATACAGCATTCATAGTAGAAGAGAGACTTTTGGTGGAGTCAC 226
QY 765 TAATGGAAACCCGGAAGATAAATTAGCATTCGGGGCTCCAAACACTTGTGAAGTCCATTT 824
DB 227 TAATGGAAACCCGGAAGATAAATTAGCATTCGGGGCTCCAAACACTTGTGAAGACCATTT 286
QY 825 TGAACACACCAAGATACCTCTGTGAAACATCTTGGAGAGCTCGGAGATGGGTTTAAGGT 884
DB 287 TGAATACACCAAGATACCTCTGTGAAACATCTTGGAGAGCGCGGAGATGGGTTTAAGGT 346
```

Qy	885	GGCCATGAACATCTCTCAACAGGGCCGTTTCAGCATGGGAGGCTCGTGGCTGGCTGCT	944
Db	347	GGTCATGAACATCTCTCAACAGGGCCGTTTCAGCATGGGAGGCTCGTGGCTGGCTGCT	406
Qy	945	CAAGAGATTGATTGAAATGACTGTGACTACGCTGCACAAAGAA	989
Db	407	CAAGAGATTGATTGAAATGACTGTGACTACGCTGCACAAAGAA	451

RESULT 8

```

US-09-917-800A-1584
: Sequence 1584, Application US/09917800A
: Patent No. US20020119462A1
: GENERAL INFORMATION:
: APPLICANT: Mendrick, Donna
: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Castle, Arthur
: APPLICANT: Elashoff, Michael
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Molecular Toxicology Modeling
: FILE REFERENCE: 44921-5038-US
: CURRENT APPLICATION NUMBER: US/09/917,800A
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/222,040
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: US 60/222,880
: PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: US 60/290,029
: PRIOR FILING DATE: 2001-05-11
: PRIOR APPLICATION NUMBER: US 60/290,645
: PRIOR FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: US 60/292,336
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: US 60/295,798
: PRIOR FILING DATE: 2001-06-06
: PRIOR APPLICATION NUMBER: US 60/297,457
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,884
: PRIOR FILING DATE: 2001-06-19
: PRIOR APPLICATION NUMBER: US 60/303,459
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 1740
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1584
: LENGTH: 2117
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020119462A1 NW_012891
US-09-917-800A-1584

```

	Query Match	20.9%	Score 388.6;	DB 10;	Length 2117;
	Best Local Similarity	54.5%;	Pred. No. 9.e-113;		
	Matches 891;	Conservative 0;	Mismatches 729;	Indels 15;	Gaps 5;
QY	113	CTTTTCGCGCAAGAGCTTTTCTCTAGGCAAAATCAAGNAAGAAAGTATTTTCCCATTTCCAG	172		
Db	236	TTTTTGTGTGGNATGTTCAAGGGCCAGCTTACCACCGACAGGTGTTCATACCCAT	295		
QY	173	AGTTTAGCCAAAGATGAACTTAATGAA---ATCAATCAGTTCCTTGGGACCCGTGGAAAAAT	229		
Db	296	CTGTGCTCAATGAAGGACAGACACAATTTCTCAAAGAGCTGTGTGGGACAGCTGGCCCGGT	355		
QY	230	TCTTTCATCGAAGAGGTGGACTCCGAAAAAATTTGACCGAAGGGAANAATCCCAATGAAA	289		
Db	356	TCSTTGAAGAAGTGAATGATCCCTGCCAAGAATGACTCTCTGGAAAGGTGGAGAGGACA	415		
QY	290	CTTTTGGAAATTTGAAGACCTTAGGCTTTTTTGGCTGCAAGTCCCAAGAAATATGCGT	349		
Db	416	CTTTTCAGGGGACTCAAAGAAACCTGGGGGCATTTGGTCTGCAAGTACCCAGCAGCTGGGTG	475		

QY	350	GCCTGGCCTTC	CAACACCAATGTACTCAAGACTAGGGAGATCATCAGCATGGATG---	400
Db	476	GTTTGGGCGCTCT	TAATACCCAGTACGGCTCGCTTGGCAGAGATTCTGGGCATGCATGACC	535
QY	407	GGTCCATCACTGT	GACCCCTGGCAGCGCACAGGCTATTGGCCTCAAGGGGATCATCTTGG	466
Db	536	TTGGGTGTAGCGTT	ACCCCTGGGAGCCCATCAGAGCATCGGTTTCAAGGGCATCTTGGCTCT	595
QY	467	CTGGCACTGAGAG	CAGAAAGCCAAATACTTTGCCTTAAACTTGGCGTCCGGGAGCACAATTG	526
Db	596	ATGGACAAAGGCC	CAGAAAGAAATAACCTCCCCAGAGTGGCATCCGGCAGGCTTTGG	655
QY	527	CAGCCTTCTGCCT	CACGGAGCCAGCTGGGAGCGATGCAGGCCTCAATCCGGAGCAGAG	586
Db	656	CGGCTTTTGCCT	CTGACTGAGCCCTCGAGGGGTCCGATGTGGCCTTATTAAGAACTCAG	715
QY	587	CCACACTAAGTGA	AGACAAAGCACTACATCCTCAATGGCTCCAAGTCTCTGGATTACTA	646
Db	716	CTGTACCTAGCCC	TGTGGAAAGTATTACTCTCAACGGAAGCAAGATCTGGATCAGTA	775
QY	647	ATGGAGGACTGCC	CAATATTTTACTGTGTTTGCNAAGACTGAGGTCTGTGATCTGTG---	703
Db	776	ATGGGGGTCTGC	GACATTTTCACTGTCTTTGGCCAAACGCCAATTAAGATCAGGCCA	835
QY	704	ATGGATCAGTGA	AGACAAAAATCACAGCATTCATAGTAAGAGAGACTTTGGTGGAGTCA	763
Db	836	CGGGGGCGGTGA	AGAGAAGATCACAGCTTTCGTAGTGAACGGAGCTTTGGAGGGGTTA	895
QY	764	CTAATGGAAACCC	GAAGATAAATTTAGGCATTCGGGGCTCCAACACTTGTGAAGTCCATT	823
Db	896	CCCATGGGCTCCC	CGGAAAGAGATGGGCATCAAGGCATCTTAACACATCAGAGGTGACT	955
QY	824	TTGAAACACCAAG	ATACCTCTGTGAAAACATCCTTTGGAGAGTCSGAGATGGGTTTAAGG	883
Db	956	TTGATGGAGTCA	AGGTGCCACAGAGATGTCTAGGAGAAGTGGGAGATGGCTTTCAAGG	1015
QY	884	TGGCCATGAACAT	CTCTCAACAGCCGGCGTTTTCAGCATTGGCAGCGTCTGGCTGGGGTGC	943
Db	1016	TTGTGTCAACAT	CTCTCAACACGGAAGATTTGGGATGGCTGCAACCTAGCAGGCACCA	1075
QY	944	TCAAGAGATTCA	TGTAATGACTCTCGTAGTACGCTGCACNAGGAACAGTTTANCAAGA	1003
Db	1076	TGAAGCCATCAT	TGGCCAGCGGTTTGATCATGTCTAACCCTTACCAGCTTTGGGGACA	1135
QY	1004	GGCTCAGTGAAT	TTGGATTGATTCAGGAGAAATTTGCACATGATGGCTCAGAAGCTTTACG	1063
Db	1136	AAATTCACAACT	TTTGGGTGATCCAGGAAAGCTGGCTCGGATGGCTATTCTCGAGTATG	1195
QY	1064	TCATGGAGAGTAT	GACCTACCTCACACGAGGATGCTGGACCAACCTTGGCTTTCCCGACT	1123
Db	1196	TGACTGAGTCCAT	GGCTTTAGCTGTAGTGCACACATGSGACCAAG---GGATTTCAAAGACT	1252
QY	1124	GCTCCATCGAGGC	AGCCATGGTGAAGGTGTTTCAGCTCCGAGGCGCTCGCAGTGTGTGA	1183
Db	1253	TCCAGATAGAAG	CTGCCATCAGCAAAATCTTTTGGCTCGGAGCGGCTCGAAAGTGCAG	1312
QY	1184	GTGAGGCGCTGC	AGATCCTCTGGGGCTTGGGCTACACAAGGGACTATCCCTCAGACGGCA	1243
Db	1313	ATGATGCATCCAG	ATTAATGGGGGCAATGGGCTTCATGAAGAACCAACGAGGCTTAGAGCGTG	1372
QY	1244	TACTGCGGTGAC	CCCCGATCCTCTCATCTTTTCAGGGAACCAATGAGATTCTCCGGATGT	1303
Db	1373	TGCTCCGAGATAT	TCGNATCTTCCGGATCTTTGAGGGACAAATGACATTTCTTCGACTGT	1432
QY	1304	ACATCGCCCTTG	ACGGGTCTGCAGATGCGCGCGCATCCTGTACTACAGGATCCATGAGC	1363
Db	1433	TTGTGGCTCTACA	AGGCTGCATGGACAAAGGAAGAACTCACGGGACTTTGGTTAATGCC	1492
QY	1364	TTAAACAGGCCAA	AGTAGTCAGGCACAGTCATGGATACCGTTTGGCCGGAGGCTTCGGGACTCCC	1423
Db	1493	TAAAGAACTCTCT	TGGANAATTTGGGCTCTCTATGAGAGAAGCAACCAACAGCTGAGGC	1552
QY	1424	TGGGCCGAACTGT	GGACCTTGGAGCTGACAGGCCAACCATGGAGTTGTGTGCACCCCACTTTTG	1483


```
Db 1553 GCGGACAGGGATTGGCAGTGGTCTGA---GTCTCTCGGAATTGTCCACCCAGAGTTGA 1609
QY 1484 CGGACAGTGCCCAACAGTTTGGAGAGAACACCTACTGCTTGGCCGCGAGCGTGGAGACAC 1543
Db 1610 GTCCGAGTGGTGAACCTGGCAGTGCAGGCTCTGGAACAATTTGCCACTGTAGTGGAGGCGA 1669
QY 1544 TGCTGCTCCGCTTTGGCAGACCATCATGGAGGAGCAGCTGTACTGAAGCGGGTGGCCA 1603
Db 1670 AGCTGATGAAGCACAGAAGAGGATGTCAATGAACAGTTCTCTGTGAGGACTGGCAG 1729
QY 1604 ACATCCTCATCAACCTGTATGACGAGCGCGTGTCTGCGGGCCAGCCGCTCCATCC 1663
Db 1730 ATGGAGCCATTGACCTCTACGCCATGGTGGTGTCTCTCCAGAGCCTCAAGATCCCTGA 1789
QY 1664 GCATTGGGCTCGCAACACGACGAGGAGTCTCTTGGCCAAACACCTTCTCGGTGGAG 1723
Db 1790 GTGAGGGCTACCCGACAGCAGCATGAGAAAATGCTGTGATAGTTGGTGCATTGAGG 1849
QY 1724 CTTACTTGCAGATC 1738
Db 1850 CTGCAACAGGATTC 1864
```

RESULT 9

```
US-09-918-995-16927
; Sequence 16927, Application US/0918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR FILING DATE: 2001-07-30
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16927
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-16927
```

```
Query Match 14.08; Score 261.4; DB 11; Length 490;
Best Local Similarity 99.3%; Pred. No. 1.6e-72;
Matches 273; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1560 CAAGCATTATGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACATCTCATCAACCT 1619
Db 217 CCAGACCATCATGAGGAGCAGCTGGTACTGAAGCGGGTGGCCAAACATCTCATCAACCT 276
QY 1620 GTATGGCATGACGGCGGTGCTGCGGGGCCAGCCGCTCCATCCGATTGGGTCCGCAA 1679
Db 277 GTATGGCATGACGGCGGTGCTGCGGGGCCAGCCGCTCCATCCGATTGGGTCCGCAA 336
QY 1680 CCAGGACGAGGTTCTCTTGGCCAAACCTTCTCGGTGGAAGCTTACTTGCAGATCT 1739
Db 337 CCAGGACGAGGTTCTCTTGGCCAAACCTTCTCGGTGGAAGCTTACTTGCAGATCT 396
QY 1740 CTTGAGCTCTCTCAGCTGACAGATGATGCTCCAGAAAACCTAGATGAGCAGATTAAGAA 1799
Db 397 CTTGAGCTCTCTCAGCTGACAGATGATGCTCCAGAAAACCTA-ATGAGCAGATTAAGAA 455
QY 1800 AGTGTCCAGCAGATCCTTGAGAGCGAGGCTATA 1834
Db 456 AGTGTCCAGCAGATCCTTGAGAGCGAGGCTATA 490
```

RESULT 10

```
US-09-822-849A-480
; Sequence 480, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 480
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-480
```

```
Query Match 11.5%; Score 214.6; DB 9; Length 312;
Best Local Similarity 96.1%; Pred. No. 1e-57;
Matches 220; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGAGCGGCTGCGGGCTCTTCTCGCACACCGGCTGCGCTGCGCTGCGGGGTCTG 60
Db 57 ATGAGCGGCTGCGGGCTCTTCTCGCACACCGGCTGCGCTGCGGGGTCTG 116
QY 61 GTGGTCTCTACCGCGAACCAGCGGCTACTGCGCACCGCCCTGTAGAGCTTTGCGC 120
Db 117 GTGGTCTCTACCGCGAACCAGCGGCTACTGCGCACCGCCCTGTAGAGCTTTGCGC 176
QY 121 AAAGAGCTTTTCTTAGGCAAAATCAAGAAAGAGTTTCCCATTTCCAGAAAGTTAGC 180
Db 177 AAAGAGCTTTTCTTAGGCAAAATCAAGAAAGAGTTTCCCATTTCCAGAAAGTTAGC 236
QY 181 CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGACCCGCTGAAAAAT 229
Db 237 CAAGATGAACCTTAATGAATCAATCAGTTCTTGGGCTAACTCTCTGGAAT 285
```

RESULT 11

```
US-09-796-692-8175
; Sequence 8175, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
```



```

; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8175
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8175

```

```

; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8175
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8175

Query Match      8.9%; Score 165.2; DB 14; Length 568;
Best Local Similarity 58.0%; Pred. No. 9.2e-42;
Matches 312; Conservative 0; Mismatches 223; Indels 3; Gaps 1;

QY      782 ATAAATAGGCATTCGGGCTCCAAACACTGTGTGAAGTCATTTTGGAAACACCAAGATAC 841
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      3 AGAAGATGGGCATCAAGGCTTCAAACACAGCAGAGGTGTCTTTGATGGAGTACGGGTGC 62

QY      842 CTGTGGAACAACATCCCTTGGAGAGTCCGAGATGGGTTTAAAGTGGGCATGGAACATCCTCA 901
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      63 CATCGGAGAACCTGTCTGGGTGAGTTGGGAGTGGCTTCAAGGTTGCCATGCATCCTCA 122

QY      902 ACAGCGCGCGGTTTCAGCATGGCAGCGTCGTGGCTGGGCTGCTCAAGAGATTGATTGAAA 961
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      123 AACATGGAAGGTTTGGCATGTGCTGCGCGCCTGGCAGGTACCATGAGAGGCATCATTCGTA 182

QY      962 TGACTGCTGAGTACGCCCTCCACAAGGAACAGTTTTAAAGAGGCTCAGTGAATTTGGAT 1021
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      183 AGCGGTAGATCATGCCACTAATCGTACCAGGTTTGGGGAGAAATTCACAACCTTTGGGC 242

QY      1022 TGATTTCAGGAGAAATTTGCACATCGATGGCTCAGAAGGCTTTACGTCATGGAGAGTATGACCT 1081
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

1082	ACGTCACAGCAGGGATGCTGGACCAACCTTGGCTTTCCGACATGCCTCATCGAGGAGGCCA	1141
QY		
Db		
303	ACATGGTGAGTGCTACATGGACCA---GGGAGCACCGGACATTCAGATAGAGGGCCGCA	359
QY		
1142	TGGTGAAGTGTTTCACTCCGAGGCGCGCTGGCAGTGTGTGAGTGGGCGCTGCAGATCC	1201
QY		
Db		
360	TCAGCAAAATCTTTGGCTCGGAGGCGCCTGGAAGGTGACAGTGAATGCATCCAAATCA	419
QY		
1202	TCGGGGGCTTGGCTTACACAAGGGGACTATCCCTACGAGGGCATCTGCGGTGACACCCGCA	1261
QY		

Db 420 TGGGGGGTATGGCTTCATGAAGAACCTGGAGTAGAGCGTGTGCTCCGAGATCTTCGCA 479
QY 1262 TCCTCTCATCTTCGAGGAACAATGAGATCTCCGGATGATACATCGCCCTGACGGG 1319
Db 480 TCTTCGGGATCTTGAGGGGCAATGACATCTTCGGCTGTGTTGGGCTCTGCAGGG 537

RESULT 13

US-09-867-701-3392
; Sequence 3392, Application US/09867701
; Patent No. US2002013237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3392
; LENGTH: 183
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(183)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-3392

Query Match 8.6%; Score 159.4; DB 10; Length 183;
Best Local Similarity 97.7%; Pred. No. 3.2e-40;
Matches 171; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1271 TCTTCAGGGAACCAATGAGATCTCCGGATGATACATCGCCCTGACGGGTCTGCAGCATG 1330
Db 9 TCTTCAGGGAACCAATGAGATCTCCGGATGATACATCGCCCTGACGGGTCTGCAGCATG 68
QY 1331 CGGGCCGATCCTGACTACAGGATCCATGAGCTTAAACAGGCAAGTGAGCACAGTCA 1390
Db 69 CGGGCCGATCCTGACTACAGGATCCATGAGCTTAAACAGGCAAGTGAGCACAGTCA 128
QY 1391 TGATACCTGTGGCCGGAGG-CTTCGGGACTCCCTGGGCGCAACTGTGACCTGG 1444
Db 129 TGGATACCTGTGGCCGGAGGTTTCGGGACTCCCTGGGCGCAACTGTGACCTGG 183

RESULT 14

US-10-156-761-5259
; Sequence 5259, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 5259
; LENGTH: 1158
; TYPE: DNA

; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1158)
US-10-156-761-5259
Query Match 8.1%; Score 151.2; DB 14; Length 1158;
Best Local Similarity 49.3%; Pred. No. 4.3e-37;
Matches 511; Conservative 0; Mismatches 513; Indels 12; Gaps 4;
QY 278 TCCAGATCAAACTTTGGAGAAATTCAGAGAGCTAGGGCTTTTGGGCTCGCAAGTCCCAG 337
Db 110 TCCGTACGAGATCGTCCGCGAGATGGCGGATGGGCTGTTCGGGCTGCGGTTCCCGG 169
QY 338 AAGATATGTGGCTGGCTTCTCCAA---CACCATGTACTCAAGACTAGGGGAGATCA 394
Db 170 AGGAGTACGGCGCATGGCGGCGACTATCTGCGCTCGGCATCGCCCTCGAAGAACTCG 229
QY 395 TCAGCATGATGGTCCATCATCTGACCTGCGCAGCGCACAGGCTATTTGGCTCAAG 454
Db 230 CCGCGTCTGACTCTCTCGCTCGCATCACCTGGAGCGCGGGTCTCTACTGGCGCGATGC 289
QY 455 GGATCATCTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTGCCTAAACTGGCTCCG 514
Db 290 CGATCCACTCTTCGGGACGGCGCGAAGAGCGGAGTGGCTCCCGCGCTGTGTCCG 349
QY 515 GGGAGCACATTCAGGCTTCTGCTCCACGGAGCCAGCCAGTGGGAGCGGATGACAGCTCAA 574
Db 350 GCGAGATCTCGCGGCTTCTCGGTCTCACGAGCGGCGGCTCGGAGCGCGGCGCA 409
QY 575 TCCGGAGCAGCCACACTAAGTGAAGACAAGACACTATACCTCAATGGCTCCCAAGG 634
Db 410 CCGCAGCAGCGCCCGCTGGAGTGCAGCAACGAATGGGTGATCAACGGCACCAGT 459
QY 635 TCTGGATTACTAATGAGGACTGGCCAATATTTTACTGTGTGTTGCAAGACTGAGGTG 694
Db 470 GCTTCATCACTACTCGGCA---CCGACATCAGCGGTTGGTGAGGTCACGCGGTCA 526
QY 695 TTGATTCTGATGGATCAGTGAAAGACAAATCACAGCATTCATAGTAGAAGAGACTTTG 754
Db 527 CCGCGCGAAGCCCGACGAGAAACCGCTGATCTCTCGATCATCTGCTCCGCTCGGACGC 586
QY 755 GTGGAGTCTAATGGAACCCGAAAGATAATTAGGTCATTCGGGCTCCCAACTTTGTG 814
Db 587 CCGGCTTCACGTCGCGCCCGCTACTCGAAGTTCGCTGGAACGCTCGGACACCCGCG 646
QY 815 AAGTCCATTTTGAAGAACACCAAGATACCTGTGGAAGAACATCTCTTGGAGAGTGGAGATG 874
Db 647 AGCTGTCTTCGCGGAGTCCGCGTCCCGCGCGCAACCTGCTGGGCGAAGACAGGCGCG 706
QY 875 GGTTTAAGGTGGCCATGAACATCTCAACAGCGCGCGTTTACAGATGGGCGGCGTGTGG 934
Db 707 GGTACGCGAGTTCTTCGCGATCTTCGAGGAGGAGGATCGCCATCTCGGCCCTCGCCA 766
QY 935 CTGGGCTCTCAAGAGATTGATTAATGACTCTGAGTACGCTGCACAGGAAGAAACAGT 994
Db 767 CCGGCTTCGCGGAGGCTGGTGAGGAGTGGTGAAGTACGCGGGGGAACGCGACCGGT 826
QY 995 TTAACAAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGACATGATGGCTCAGA 1054
Db 827 TCGGCGGAACATCGGGGCGTACCAGGCCATCCAGTTCGAAGATCGCGGATATGAGATGA 886
QY 1055 AGGCTTACGTGATGAGAGATGATGACCTACTCACAGCGGATGCTGGACCACTGGCT 1114
Db 887 AGCGCA---CATGGCCCGCTGGCTGGCTGACCGCGCTCCCGCTGGTCCGCGCGG 943
QY 1115 TTCCCGACTGCTCCATCGAGGAGCGCATGGTGAAGTGTTCAGCTCCGAGGCGCGCTGGC 1174
Db 944 AGCC---CTTCAAGAGAGGCGCGCATCGCCAGCTCTACTCTCCAGGTCGCGCTCG 1000
QY 1175 AGTGTGTGATGAGCGCTGAGATCCTCGGGGCTTGGGCTACACAAGGAGACTATCCGT 1234
Db 1001 ACAACGCCCGGAGGCGCCAGATCCACGCGGCTACGGCTTTCATGACGAGTACCGCG 1060

Qy	1235	ACGAGCGCATACGCGTGACACCGGATCCTCTCATCTTCGAGGAAACCAATGAGATTC	1294
Db	1061	TGCCCGGATGTGGCGGACTCCAAGATCCTGGATCGCGAGGCACCGCGAGTGTC	1120
Qy	1295	TCCGGATGTACATCGC	1310
Db	1121	ACGCGATGCTGATCGC	1136

RESULT 15

```

US-10-156-761-1/c
: Sequence 1, Application US/10156761
: Publication No. US20030119019A1
: GENERAL INFORMATION:
: APPLICANT: OMURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, MASAHIRA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-262
: CURRENT APPLICATION NUMBER: US/10/156.761
: CURRENT FILING DATE: 2002-05-29
: PRIOR APPLICATION NUMBER: JP 2001-204089
: PRIOR FILING DATE: 2001-05-30
: PRIOR APPLICATION NUMBER: JP 2001-272697
: PRIOR FILING DATE: 2001-08-02
: NUMBER OF SEQ ID NOS: 15109
: SEQ ID NO 1
: LENGTH: 9025608
: TYPE: DNA
: ORGANISM: Streptomyces avermitilis
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (4187715)
: OTHER INFORMATION: a, t, c, g, other or
US-10-156-761-1

```

Query Match	8.1%;	Score 151.2;	DB 14;	Length 9025608;
Best Local Similarity	49.3%;	Pred. No. 1.4e-34;		
Matches 511;	Conservative 0;	Mismatches 513;	Indels 12;	Gaps 4;
QY	278	TCCAGATCAAACTTTGGAGAAATCGAAGAGCCTAGGCTTTTGGCTGCCAAGTCCAG	337	
Db	6391537	TCCCGTACGAGATCGTCCCGGAGATGGCGCGATGGCCCTGTTGGGCTGCCCTTCCCG	6391478	
QY	338	AGATATAGTGGCTGGCTTCTCCAA---CACCATGTACTCAAGACTAGGGAGATCA	394	
Db	6391477	AGAGTACGGCGGCATGGCGCGGACTATCTGCGCTGGCATCGCCCTCGAAGAAGTCTG	6391418	
QY	395	TCAGATGATGGTGGTCCATCACTGTGACCCCTGGCAGCGCACAGGCTATTGGCCTCAAGG	454	
Db	6391417	CCCGGTGACTCTCCGTGCGCATCAACCTGGAGCCGGGGTCTCAGTGGGCGCATGC	6391358	
QY	455	GGATCATTTGGCTGGCACTGAGGAGCAGAAAGCCAAATACTTGCCTAAACTGGCGTCCG	514	
Db	6391357	CGATCCACTCTTCGGAGCGGACGCGCAGAAGCGGAGTGGCTCCCCCGGCTGTGTTCCG	6391298	
QY	515	GGAGCACATTCACGGCTTCCCTTCACGGAGCCAGCAGTGGGAGCGATGCAGCCTCAA	574	
Db	6391297	CGGAGTCTCGGGCCCTTCGGTCTCACGAGCGCGACCGCGGCTCGAGCGGGGGCGCA	6391238	
QY	575	TCCGGAGCAGAGCCACACTAAGTGAAGACAAGAAGCACTACATCCTCAATGGCTCCAAAGS	634	
Db	6391237	CGCGCACAGCGCCCGCTTGGACGAGTGCAGCAACGAATGGTGATCAACGGACCAAGT	6391178	
QY	635	TCTGGATTACTAATGGAGACTGGCCAATATTTTTACTGTGTTTGCAAAGACTGAGGTCG	694	
Db	6391177	GCTTCATCACCACATCGGGCA---CCGACATCACGGGGTGTGTGACGGTTCACGGGGGTCA	6391121	

Qy	695	TTGATTCGTGATGATCAGTGAAGACAAAATCACAGCAATTCATAGTAGAAGAGACTTTG	754
Db	6391120	CCGGCCGGAAGCCGACGGGAACCGGTGATCTCCTCGATCACTGCTCCGTCGGCACGC	6391061
Qy	755	GTGGAGTCACCTAATGGGAAACCCGAAGATAAATAGGCAATCGGGGCTCCAACACTTGTC	814
Db	6391060	CCGGCTTACCGGTGCGCGCCCGCTACTCGAAGGTTCGGCTGGAACGCTCGGACACCCGCG	6391001
Qy	815	AAGTCCATTTTGAACACCAACAGATACCTGTGGAAACATCTTTGGAGAGGTCGGAGATG	874
Db	6391000	AGCTGTCTTCGCGGACGTCCGGTCCCGGCCGGAACCTGCTGGCGGAACAGGGCCGCG	6390941
Qy	875	GGTTTAAGGTGGCCATGAACATCCTCAACAGCGCCGGTTTCAGCATGGGCAGCGTCGTGG	934
Db	6390940	GGTACGCGCAGTTCTTCGCGATCCTCGACAGGACGGATCGCCATCTCGGCCCTGGCCA	6390881
Qy	935	CTGGGCTGCTCAAGAGATGATTGAATGACTGCTGAGTACGGCTGCACNAGAAACAGT	994
Db	6390880	CCGGGCTCGCCAGGGCTCGGTGGAGAGTCGGTGAAGTACGCCGGGGACGGCACGCGT	6390821
Qy	995	TTAACAGAGGCTCAGTGAATTTGGATTGATTCAGGAGAAATTTGCACGTGATGGCTCAGA	1054
Db	6390820	TCGGGGCGAATACCGGGGGCTACAGGCCATCCAGTTCAAGATCGCCGATATGGAGATGA	6390761
Qy	1055	AGGCTTACGTCATGAGAGATGACCTACTCACACGAGGATGCTGGACCAACCTGGCT	1114
Db	6390760	AGCGCA--CATGGCCCGGTGGGTGGGTGACCCCGGTCCCGGCTGCTGCCGGGG	6390704
Qy	1115	TTCCGAGCTGCTCCATCGAGGACGCCATGGTGAAGGTGTCAGCTCCGAGGCCGCTGGC	1174
Db	6390703	AGGC--CTTCAAGAAAGAGGGCGGCATGCCAAGCTCTACTGCTCCAGGGTCGCGGTGC	6390647
Qy	1175	AGTGTGTGATGAGGCGCTCGAGATCCTCGGGGGCTTGGGCTACACAAGGCACTATCCGT	1234
Db	6390646	ACAACGCCCGAGGCCACCCAGATCCACGGCGCTACGGCTTCATGAACGAGTACCCCG	6390587
Qy	1235	ACGAGCGCATGCGGTGACACCCCGATCCTCGTCATCTTCGAGGGAACCAATGAGATTTC	1294
Db	6390586	TGCCCCGGATGTGGCGCACTCCAAGATCCTGGATCGCGGAGGCGACACGCGAGGTGC	6390527
Qy	1295	TCGGATGTACATCGC	1310
Db	6390526	AGCGGATGCTGATCGC	6390511

Search completed: September 7, 2003, 05:59:04
Job time : 432.388 secs

THIS PAGE BLANK (USPTO)